THIS WEEK

Repeat after me

With plagiarism seemingly endemic in Romania, as well as rife among Europe's political class, a bid by academics to root out misconduct deserves widespread support.

Elena Ceauşescu did not have a BSc, but the power of her husband Nicolae — Romania's dictator until communism fell — still made sure that the University of Bucharest awarded her a PhD in chemistry. The contents of her many scientific papers were penned by others.

The couple were executed on Christmas Day 1989 for crimes more terrible than poor publication ethics. But their practice of playing fast and loose with academic principles has flourished in notoriously corrupt post-communist Romania.

Over the past 18 months, *Nature* has chronicled an epidemic of plagiarism involving prominent political figures in Europe, reporting, among others, on Germany's former defence minister Karl-Theodor zu Guttenberg, who plagiarized material for his law thesis; Hungary's former president Pál Schmitt, whose thesis on physical education contained plagiarized material; as well as Romania's prime minister Victor Ponta, who continues to dismiss well-founded accusations that he plagiarized sources for his law thesis. On page 264, we highlight widespread plagiarism within Romania's universities and the worrying fact that so many academics there seem not to realize why this is a problem.

Plagiarism seems to be disturbingly prevalent among the European political class and in Romanian academia, but cases continue to pop up everywhere. The Internet makes it easier to detect, but also easier to perpetrate, because anyone can cut and paste more or less anything in a matter of seconds. Perhaps it is this easy access to the words of others that encourages some academics to think that plagiarism is not a serious issue.

They are wrong. Plagiarism is illegal (theft of intellectual property) and immoral, and anyone whose reputation and career rely on publishing their ideas and findings needs to care about it. True, it does not directly affect the scientific literature in the same way as other types of misconduct such as data fabrication. But it has an indirect impact on the academic system because it helps to promote the careers of the fraudulent and the undeserving. And those who climb the academic ladder on the back of dishonest publication records often imbue their students with the same disrespect for scientific method and academic principles.

It matters on a larger scale, too. Most countries accept that to attain economic prosperity they need a robust research base, a concept enshrined in the European Union's Treaty of Lisbon. But a research base contaminated with plagiarism can never function optimally. Romania, a signatory to the treaty, seemed to be on track to a more honest and promising future when it passed its education law, designed to inject competition into its universities and root out widespread scientific misconduct. Yet that law is now being undermined by political interference in the very ethics councils that should be helping to implement it. Against this backdrop, it is easy to see why Romania's excellent scientists — and there are many of them — choose to work mostly outside the country.

A group of researchers is now trying to change things, partly through a website to track and investigate cases of misconduct in Romania. Their stated goal is to "reform and restore confidence" in the country's academic system. Scientists everywhere should back their effort and pass on their message — with appropriate attribution, of course. ∎

Safety shambles

Lax management of Fukushima clean-up intensifies concerns over Japan's nuclear future.

Earlier this summer, it emerged that five people working to clean up Japan's devastated Fukushima nuclear power plant had covered their dosimeters with lead. They blocked the radiation detectors — intended to alert them when their exposure was reaching dangerous levels — late last year to allow them to work longer hours.

An announcement made last week by the Tokyo Electric Power Company (TEPCO), which runs the plant, found that the five workers were contracted by a subcontractor of a subcontractor, and had not even been authorized to work at the plant. Other workers were found who were not using the devices at all. In response, TEPCO has barred its direct subcontractor, Tokyo Energy and Systems, from bidding for contracts for three months. The ministry of health and labour is still investigating.

That such egregious flouting of safety protocols would occur despite the media attention on the clean-up efforts is astonishing. And it seems all the more so given the ongoing concern about the health risks of radiation in Japan, and that the nuclear industry's lack of transparency and cavalier attitude towards safety are fuelling a debate over atomic power that is becoming the most politically divisive in the country for decades.

That's on the ground. Decision-making looks no prettier at TEPCO headquarters. Last week, the company made available some 150 hours of video recordings of exchanges between the company and staff at its nuclear plants in the first days of the disaster. The videos highlight some uncomfortable facts that the company had previously denied — that senior officials hesitated to use sea water to cool overheating reactors for fear of damaging them, for example, despite recommendations from a local plant manager to do so.

In the days after 11 March 2011, the world saw that TEPCO's reactors were not designed to withstand disaster. Now it is becoming clear how poorly the company has dealt with the aftermath. ∎

Why we are poles apart on climate change

The problem isn't the public's reasoning capacity; it's the polluted science–communication environment that drives people apart, says **Dan Kahan**.

Understandably anxious to explain persistent controversy over climate change, the media have discovered a new culprit: the public. By piecing together bits of psychological research, many news reporters, opinion writers and bloggers have concluded that people are simply too irrational to recognize the implications of climate-change science.

This conclusion gets it half right. Studying things from a psychological angle does help to make sense of climate-change scepticism. But the true source of the problem, research suggests, is not that people are irrational. Instead, it is that their reasoning powers have become disabled by a polluted science-communication environment.

Social-science research indicates that people with different cultural values — individualists compared with egalitarians, for example — disagree sharply about how serious a threat climate change is. People with different values draw different inferences from the same evidence. Present them with a PhD scientist who is a member of the US National Academy of Sciences, for example, and they will disagree on whether he really is an 'expert', depending on whether his view matches the dominant view of their cultural group (D. M. Kahan *et al. J. Risk Res.* **14,** 147–174; 2011).

The positions on climate change of both groups track their impressions of recent weather. Yet their impressions of what the recent weather has been are polarized, too, and bear little relationship to reality (K. Goebbert *et al. Weath. Clim. Soc.* **4,** 132–144; 2012). But is this sort of cultural polarization evidence of irrationality? If it is, then how can we explain the fact that members of the lay public who are the most science literate, and the most proficient at technical reasoning, are also the most culturally polarized (D. M. Kahan *et al. Nature Clim. Change* http://dx.doi.org/10.1038/nclimate1547; 2012)?

If anything, social science suggests that citizens are culturally polarized because they are, in fact, too rational — at filtering out information that would drive a wedge between themselves and their peers.

For members of the public, being right or wrong about climate-change science will have no impact. Nothing they do as individual consumers or as individual voters will meaningfully affect the risks posed by climate change. Yet the impact of taking a position that conflicts with their cultural group could be disastrous.

Take a barber in a rural town in South Carolina. Is it a good idea for him to implore his customers to sign a petition urging Congress to take action on climate change? No. If he does, he will find himself out of a job, just as his former congressman, Bob Inglis, did when he himself proposed such action.

Positions on climate change have come to

POSITIONS ON CLIMATE CHANGE HAVE COME TO SIGNIFY THE KIND OF PERSON ONE IS.

signify the kind of person one is. People whose beliefs are at odds with those of the people with whom they share their basic cultural commitments risk being labelled as weird and obnoxious in the eyes of those on whom they depend for social and financial support.

So, if the cost of having a view of climate change that does not conform with the scientific consensus is zero, and the cost of having a view that is at odds with members of one's cultural community can be high, what is a rational person to do? In that situation, it is perfectly sensible for individuals to be guided by modes of reasoning that connect their beliefs to ones that predominate in their group. Even people of modest scientific literacy will pick up relevant cues. Those who know more and who can reason more analytically will do a still better job, even if their group is wrong on the science.

So whom should we 'blame' for the climate-change crisis? To borrow a phrase, it's the 'science-communication environment, stupid' — not stupid people.

People acquire their scientific knowledge by consulting others who share their values and whom they therefore trust and understand. Usually, this strategy works just fine. We live in a science-communication environment richly stocked with accessible, consequential facts. As a result, groups with different values routinely converge on the best evidence for, say, the value of adding fluoride to water, or the harmlessness of mobile-phone radiation. The trouble starts when this communication environment fills up with toxic partisan meanings — ones that effectively announce that 'if you are one of us, believe this; otherwise, we'll know you are one of them'. In that situation, ordinary individuals' lives will go better if their perceptions of societal risk conform with those of their group.

Yet when all citizens simultaneously follow this individually rational strategy of belief formation, their collective well-being will certainly suffer. Culturally polarized democracies are less likely to adopt polices that reflect the best available scientific evidence on matters — such as climate change — that profoundly affect their common interests.

Overcoming this dilemma requires collective strategies to protect the quality of the science-communication environment from the pollution of divisive cultural meanings. Psychology — along with anthropology, sociology, political science and economics — will play a part. But to apply the insights that social science has already given us, we will have to be smart enough to avoid reducing what we learn to catchy simplifications. ∎

Dan Kahan *is the Elizabeth K. Dollard professor of law and professor of psychology at Yale Law School in New Haven, Connecticut. e-mail: dmkahan@gmail.com*

Genetics of sexual harassment

A gene important in limb development that is normally suppressed after birth can cause female mice to become sexually aggressive, biting the genitals of males, when it is anomalously expressed in the brain.

Jozsef Zakany and Denis Duboule at the University of Geneva, Switzerland, deleted a tiny section of the mouse genome that includes a cluster of developmental genes called *HoxD*. The deletion resulted in the abnormal expression of one *Hox* gene in small populations of brain cells. This caused females in the sexually receptive phase of their reproductive cycle to hyperactively chase male mice, biting the males' genitals and severely mutilating them.

The finding could open up new approaches for investigating the molecular and cellular basis of female courtship behaviour, say the researchers.
Curr. Biol. http://dx.doi. org/10.1016/j.cub.2012.06.067 (2012)

S. BENSON-AMRAM

Curious hyenas crack puzzles

Spotted hyenas that exhibit a wider range of exploratory behaviours are better at solving problems than those that show a smaller range — just as creativity can aid problem-solving in humans.

Sarah Benson-Amram and Kay Holekamp at Michigan State University in East Lansing devised a steel puzzle box and baited it with a slab of meat (pictured). Over the course of a year, the researchers observed how 62 spotted hyenas (*Crocuta crocuta*) from two clans in Kenya interacted with the box. The team measured the hyenas' reluctance to touch the box — as a proxy for their fear of new things — and the number of tactical behaviours that they used to try to open it, such as biting, pushing or flipping, as well as their persistence. In both adult and juvenile hyenas, animals that showed a broader range of behaviours were more likely to succeed in opening the box.
Proc. R. Soc. B http://dx.doi.org/10.1098/ rspb.2012.1450 (2012)

Ramped up resolution

Colour images with an effective resolution of up to 100,000 dots per inch — the maximum possible owing to the properties of visible light — could now be a reality, thanks to researchers in Singapore.

Joel Yang at the Institute of Materials Research and Engineering in Singapore and his colleagues began with a silicon wafer and created tiny posts tens of nanometres in diameter on the surface. The posts were capped with silver disks. Depending on their size and spacing, groups of posts preferentially reflected a single colour back towards the viewer. Using their system, the researchers created colour images (**pictured**), which could be resolved at the diffraction limit of optical microscopes.

The posts could be used in security images and in optical filters or data storage, the team suggests.
Nature Nanotechnol. http:// dx.doi.org/10.1038/ nnano.2012.128 (2012)
For a longer story on this research, see go.nature.com/vwahok

Sloth inner-ear diversity

The three-toed sloth shows surprising diversity in the structure of its semicircular canals — part of the inner ear responsible for detecting motion. The sloths' lethargy and lack of agility might have reduced the need for sensitivity to motion, leading to relaxed selection pressure on this apparatus.

Semicircular-canal morphology is usually remarkably stable between members of the same species in placental mammals. Guillaume Billet at the University of Bonn, Germany, and his colleagues used high-resolution computed tomography to compare the semicircular canals of the three-toed sloth (*Bradypus*

variegatus) with those of faster-moving mammals. The structures showed greater variation in shape, size and orientation in sloths than in squirrels, moles or armadillos.
Proc. R. Soc. B http://dx.doi.org/10.1098/rspb.2012.1212 (2012)

MATERIALS

Polymers track the Sun

Solar-cell arrays catch more light when they can tilt to follow the Sun's path, but this motion tracking does not have to be driven by power-hungry machinery.

Hongrui Jiang and his colleagues at the University of Wisconsin-Madison mounted a platform of solar cells on top of elastic supporting columns. Each column is coated with films made up of carbon nanotubes embedded in liquid crystal elastomers — rubbery networks composed of ordered polymers. When the nanotubes in the columns nearest the Sun absorb sunlight, they convert the energy to heat, and this changes the orientation of the polymer chains, causing the columns to contract. This contraction tilts the solar-cell platform towards the Sun. As the Sun moves across the sky, different columns contract, changing the platform's tilt to track the Sun's motion.
Adv. Funct. Mater. http://dx.doi.org/10.1002/adfm.201202038 (2012)

ZOOLOGY

Beetles walk underwater

A terrestrial leaf beetle can walk underwater, thanks to tiny air bubbles trapped between hair-like structures on its feet.

NAT. INST. MATER. SCI.

Gastrophysa viridula (**pictured**) use these 'hairs', or setae, and the secreted fluid that covers them, to stick to leaves when on land. Naoe Hosoda at the National Institute for Materials Science in Tsukuba, Japan, and Stanislav Gorb at the University of Kiel in Germany found that when the animals walk underwater, air bubbles remain trapped between the setae. The bubbles not only provide adhesion but also de-wet the area around the beetles' feet, enabling the setae to function as they do on land.

On the basis of these principles, the researchers were able to develop an artificial polymer that successfully kept small objects attached to vertical surfaces underwater.
Proc. R. Soc. B http://dx.doi.org/10.1098/rspb.2012.1297 (2012)
For a video linked to this research, see go.nature.com/ecyztv

REGENERATIVE MEDICINE

Nanofibres foster blood vessels

Self-assembling nanofibres can help to spur the formation of blood vessels in the hearts of rats and pigs during recovery from experimentally induced heart attacks.

Researchers led by Patrick Hsieh of National Cheng Kung University Hospital in Tainan, Taiwan, injected protein fragments that self-assemble into nanofibres, and VEGF — a protein that promotes blood-vessel formation — into the animals' injured heart muscle. Whereas VEGF alone conveyed no noticeable benefits, animals that received both nanofibres and VEGF had more arteries and stronger heart performance four weeks after the simulated heart attacks. The nanofibres, which degrade slowly over time, seem to function as a scaffold that retains and recruits restorative cells.
Sci. Transl. Med. **4,** 146ra109 (2012)

COMPUTATIONAL BIOLOGY

'Whole-cell' computer model

★ HIGHLY READ on www.cell.com 9 Jul–8 Aug

By producing a mathematical model of an entire human bacterial pathogen, researchers have made predictions about its cellular behaviour. The model accounts for the functions of the bacterium's known genes and the interactions between its constituent molecules.

Built by Markus Covert at Stanford University in California and his colleagues, the model captures the life cycle of the bacterium *Mycoplasma genitalium*, which has 525 genes. It incorporates data from more than 1,900 experimental measurements and encapsulates 28 groups of cellular processes, from DNA replication to protein folding. Using their tool, the researchers predicted that, on average, there are more than 30,000 collisions between pairs of DNA-binding proteins on the organism's single chromosome per cell cycle — a number that would be difficult to deduce experimentally.

Whole-cell models could speed up biological research and even pave the way for computer-designed organisms, the researchers say.
Cell **150,** 389–401 (2012)

T. SPADY/OSTRANDER LAB

GENETICS

Heady dog genetics

Variation in a handful of genes explains why some dog breeds, such as pugs, have round, thickset heads, whereas others, such as collies, have long, narrow skulls.

Selective breeding by humans has had a strong influence on the shapes of dogs' heads, which vary greatly across the species. Elaine Ostrander at the National Human Genome Research Institute in Bethesda, Maryland, and her colleagues compared the shapes of 533 skulls belonging to 120 dog breeds (**a sample pictured**) and four subspecies of grey wolf with genetic data from many of the same species. Five genomic regions seem to set breeds with round heads apart from those with elongated heads.

Sequencing one of these regions in 11 dog breeds uncovered a mutation in a developmental gene called *BMP3* in round-headed bulldogs and Pekingese. Further analysis revealed this mutation in nearly all breeds with very short heads, including pugs, Boston terriers and Shih Tzus.
PLoS Genetics **8,** e1002849 (2012)

↻ **NATURE.COM**
For the latest research published by *Nature* visit:
www.nature.com/latestresearch
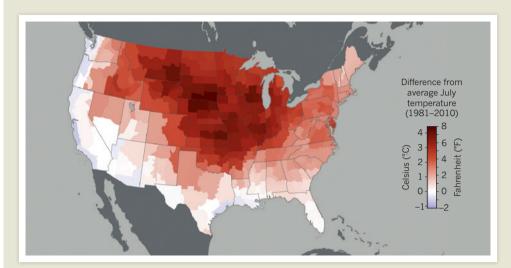
POLICY

India's GM caution

A cross-party panel of politicians in India has urged the government to halt all open-field trials of transgenic crops until it develops a better monitoring and oversight system. In a report released on 9 August, the Parliamentary Standing Committee on Agriculture in India also called for a complete overhaul of the regulatory system for transgenic-crop approvals and recommended that all genetically modified (GM) products be labelled. India has approved only one GM crop (cotton) for cultivation; GM brinjal (aubergine) was approved in 2009, but in 2010 the government banned its cultivation indefinitely, after public opposition. See go.nature.com/r6zcy8 for more.

Development work

The United Nations said on 9 August that it is to set up a global network of research centres, universities and institutes to help identify the best paths to sustainable development. The Sustainable Development Solutions Network will be directed by economist Jeffrey Sachs, who founded the Millennium Villages international development project and heads the Earth Institute at Columbia University in New York City. The move follows on from the UN's Rio+20 conference on sustainable development in June.

Swine-flu alert

The number of reported cases in an outbreak of H3N2v virus — a variant strain of swine flu that can pass from pigs to humans — took a sudden spike last week. The US Centers for Disease Control and Prevention in Atlanta, Georgia, reported 154 cases

US swelters in hottest month ever

July was the hottest month ever recorded in the United States, the National Oceanic and Atmospheric Administration revealed on 8 August. The average temperature across the continental states was 25.3 °C, one-tenth of a degree above the previous 1936 July record.

Compared with the average for 1981–2010, some parts of the country recorded lower temperatures (blue), whereas parts of the Midwest and the northern Rocky Mountains chalked up readings as high as 4.4 °C above (dark red). See go.nature.com/g94jrm for more.

on 10 August, up from 16 on 3 August; by 13 August the total had reached 176. Most are in Indiana and all are the result of contact with pigs by farmers or through agricultural fairs. Symptoms are mild, and the virus does not seem to transmit between humans. But it is raising eyebrows because it contains a gene from the H1N1 pandemic strain that may increase transmissibility among humans. See go.nature.com/z75z8k for more.

Russia's space woes

Failures in Russia's space launches have drawn the ire of the country's prime minister. Dmitry Medvedev said on 9 August that the botched launch of two communications satellites earlier that week was costing the country money and prestige, and that he would chair a meeting to look

into the failures. Russia lost three navigation satellites in December 2010; a military satellite in February 2011; a telecommunications satellite that August; and the Phobos-Grunt mission to the Martian moon Phobos in November.

EVENTS

Agent Orange

The United States launched an effort on 9 August to clean up dioxin-contaminated soil left over from the Vietnam War at a former American airbase near Da Nang city, Vietnam. The site was one of many hotspots that still have especially high concentrations of dioxins after being sprayed with herbicides such as the infamous Agent Orange (used to defoliate jungle vegetation) from 1962 to 1971. The remediation project, funded by the US government

through the US Agency for International Development, is expected to cost US$43 million and take four years.

Iran earthquakes

Two earthquakes of magnitudes 6.3 and 6.4 struck near Ahar in northwestern Iran on 11 August, injuring thousands in rural villages and resulting in a death toll reported at more than 300 people as *Nature* went to press. The region has experienced seven quakes of magnitude 6 or above in the past 40 years, the US Geological Survey says.

BUSINESS

Battery rescue

A123 Systems, a leading US manufacturer of lithium-ion batteries for electric vehicles, may have found a solution to its financial woes. The firm, based in Waltham, Massachusetts,

said on 8 August that it had reached preliminary agreement for an investment of up to US$450 million from Wanxiang Group, China's largest maker of automobile parts, based in Zhejiang. Formed in 2001, A123 is a spin-off from work at the Massachusetts Institute of Technology, Cambridge. In May, it warned investors that it was struggling for cash, and it has now reported net losses of $208 million for the first half of 2012.

Material gains

Oil and gas giant BP will invest US$100 million over ten years in a university-based research centre for advanced materials, it announced on 7 August. With a hub at the University of Manchester, UK, the centre will have smaller UK branches at Imperial College London and the University of Cambridge, and a US branch at the University of Illinois at Urbana-Champaign. It will look at research issues important to BP, such as developing metal alloys that can work in deep-water drilling and membranes for purifying oil, gas, water and biofuels. See go.nature.com/lyzvfe for more.

Drug-discovery split

Irish pharmaceutical company Elan is to split into two, completely separating its drug-discovery work from the main business. The firm, which is based in Dublin, plans to spin off its research work into a small separate company called Neotope Biosciences, which will have its operations base in South San Francisco, California, and will be given at least US$120 million in start-up funding. Elan itself will keep ownership of assets such as its stake in multiple-sclerosis drug Tysabri (natalizumab). The separation, announced on 13 August, should be complete by the end of the year, Elan said.

PEOPLE

Misconduct fall-out

A prominent Danish neuroscientist could lose her PhD and medical-sciences doctorate, after an 8 August report from a panel assembled by University of Copenhagen found evidence that she might have manipulated data in 15 papers. Milena Penkowa (**pictured**), who resigned from the university in December 2010, has denied the accusations in a response appended to the report. The government-run Danish Committees on Scientific Dishonesty will now determine whether Penkowa committed research misconduct, and an academic council at the university will decide whether she can keep

the advanced degrees she earned there. See go.nature. com/adhb1u for more.

RESEARCH

Temperature record

Physicists at the Large Hadron Collider at CERN, Europe's particle-physics laboratory near Geneva, Switzerland, say that they have achieved the hottest man-made temperatures ever, by colliding lead ions to create a quark–gluon plasma. Presenting data on 13 August at the Quark Matter 2012 meeting in Washington DC, researchers on the ALICE heavy-ion experiment said that they had created a plasma 38% hotter than the record 4-trillion-degree-Celsius plasma achieved in 2010 by a similar experiment at Brookhaven National Laboratory in New York. The measurement has not yet been converted into a temperature, but could produce something like

COMING UP

17–23 AUGUST
NASA's Mars rover Curiosity will attempt its first drive on the red planet next week — although no firm date had been set as *Nature* went to press.
www.nasa.gov/msl

19–23 AUGUST
The American Chemical Society holds its autumn meeting in Philadelphia, Pennsylvania, focusing on materials for health and medicine.
go.nature.com/tej8lw

23–25 AUGUST
The brain's connectome — along with other ways of mapping neurons and neuronal circuits — features on the agenda of the 3rd Annual Aspen Brain Forum in Colorado.
http://aspenbrainforum.com

5.5 trillion degrees Celsius, says ALICE spokesman Paolo Giubellino. See go.nature.com/kvuq7k for more.

Trust, but verify

An effort to encourage authors of high-profile research papers to get their results replicated by independent labs was launched on 14 August. The Reproducibility Initiative will work through Science Exchange, an online marketplace based in Palo Alto, California, that connects researchers wishing to outsource their experiments with service providers. The initiative will start by accepting 40–50 studies for validation, said Science Exchange. Authors will pay for validation studies themselves, and will be able to publish the results in *PLoS ONE*. See go.nature.com/gc9i9v for more.

⟳ **NATURE.COM**
For daily news updates see:
www.nature.com/news

TREND WATCH

The US Department of Agriculture predicted on 10 August that the nation's maize (corn) yields would drop to a 17-year low in 2012–13 (see chart), after a catastrophic drought that has scorched fields and sent food prices soaring. The Food and Agriculture Organization of the United Nations has suggested that the United States suspend biofuel production from maize, and US President Barack Obama unveiled an aid package for stricken farmers on 13 August.

MAIZE PROJECTIONS PLUMMET
The US Department of Agriculture has reduced its forecasts for maize (corn) production following the worst drought in 50 years.

NEWS IN FOCUS

Protesters in Qidong, Jiangsu province, persuaded officials to scrap plans for a waste pipeline.

ENVIRONMENT

Green protests on the rise in China

Environmental groups use momentum to push for reforms.

BY NATASHA GILBERT

For years, people in China have accepted murky air, tainted waters and scarred landscapes as the unavoidable price of the country's meteoric economic growth. But public dissent over environmental issues has been growing steadily in the communist nation, and now seems to be building the foundations of a fledgling green movement.

In July alone, two separate demonstrations made international news when they turned violent. About 1,000 environmental protesters in the eastern city of Qidong, Jiangsu province, successfully persuaded officials to scrap plans for a pipeline that would have carried toxic waste from a paper-manufacturing plant. The event followed a demonstration earlier that month by several thousand residents in the city of Shifang, Sichuan province, which halted the construction of a copper–molybdenum processing plant. Locals were concerned that the facility would release toxic fumes into the air and generate slag heaps that could leach arsenic into drinking water.

These recent successes come after a slew of ever-larger and more violent green protests over the past few years, as the environmental impact of China's growth becomes harder to ignore. The Chinese public has few legal or regulatory avenues to express concern about environmental damage, forcing people to demonstrate on the streets, says Li Bo, director of Friends of Nature, one of China's first green non-governmental organizations (NGOs), set up in 1993 in Beijing. "These demonstrations are evidence of the public anger and frustration at opaque environmental management and decision-making," says Bo.

Better access to mobile phones and the Internet has enabled protesters to share information and coordinate movements more effectively. The protests also seem to be shifting from rural areas — such as when 500 villagers in eastern Zhejiang province demonstrated in 2011 outside a solar-panel factory that they claimed had discharged waste into local rivers — to cities such as Qidong and Shifang, says Elizabeth Economy, director for Asia Studies at the Council on Foreign Relations, an independent think tank in New York. As the protests have become more visible, widespread news coverage has emboldened others to demonstrate despite police crackdowns, she adds.

Unlike environmental campaigns in the West, however, China's protests rarely have external coordination from green NGOs, says Economy. But some NGOs are now beginning to use the momentum created by the protests to push for regulatory and legal reforms that will improve environmental protections, says Tianjie Ma, who leads a programme on toxic pollution for the East Asia bureau of Greenpeace, the environmental campaigning organization headquartered in Amsterdam.

In what could spark a dramatic shift in green regulation in China, Friends of Nature is suing the Yunnan Liuliang Chemical Industry company for discharging 200,000 tonnes of waste into the Pearl River in the southern province of Yunnan — the first ▶

▶ time a grass-roots group has succeeded in bringing a case against a polluter in China. Friends of Nature and other green groups say that the pollution, which contains toxic chromium(VI), may have caused incidents of cancer in local villages. Bo says that the case "is a new initiative to allow pollution victims and entities like NGOs to bring cases to court against polluters". The plaintiffs are demanding that the company pays 10 million renminbi (US$1.6 million) to a compensation fund to clean up the damage.

"There are still lots of issues to be worked out," Bo adds, "such as how to do environmental damage appraisals, and it is still not clear which parties are eligible to be plaintiffs." The court has held two meetings in the past two months to discuss the case, and Bo expects a ruling by the end of the year. "We are hoping this will set a precedent" for other NGOs and members of the public to bring similar cases, he says.

Alex Wang, an environmental lawyer at the University of California, Berkeley, says that the lawsuit signals an opportunity for civil-society groups to play a part in environmental enforcement. But "only time will tell whether this turns into precedent or remains a one-time event. Too many promising initiatives like this one have languished once the initial publicity has died down," he warns.

Green groups are also calling for improvements in the environmental impact assessments (EIAs) that companies must conduct before projects are given permission to go ahead. The public has 14 days to submit concerns and objections to proposed projects in China, but locals rarely have access to detailed plans and EIAs before it is too late, says Ailun Yang, who analyses emerging economies at the World Resources Institute, a think tank in Washington DC.

Yang says that companies should be required to disclose detailed information about the technologies to be used in projects, and to explain how they will dispose of waste. The government should also produce guidance for the public to make it easier for people to participate, she adds.

Keping Ma, a plant ecologist at the Institute of Botany, part of the Chinese Academy of Sciences in Beijing, says that China's central government is developing a "more positive attitude" towards ecological and environmental issues. And local officials did act quickly to cancel the projects in Qidong and Shifang, he points out.

But the strongest pressure for reform comes directly from members of the public, who feel that their lives are threatened by pollution and environmental damage, says Bo. "Throughout several dynasties, China has a history of villages rising up and overthrowing regimes they could not tolerate," he says. "I hope history does not repeat itself, because this has a high cost." ∎

Mars scientists await feast of data

As the Curiosity rover prepares to take its first trip across the surface, the lead scientist shares his hopes for the mission.

BY ERIC HAND

John Grotzinger folds his rangy frame into a plastic leather booth at Conrad's, an unassuming diner in Pasadena, California. It is nearly midnight, and a smattering of customers have drifted in for a late-night snack. In the booth on one side, members of a Neil Young tribute band relax after a gig; on the other, a personal-trainer-cum-actor is winding down from a script reading.

But Grotzinger, who is chief scientist for NASA's Mars Science Laboratory mission, is trying to gear up for the start of his working day at the nearby Jet Propulsion Laboratory, as the rover Curiosity begins its fourth Martian day. Still riding the high of the rover's triumphant landing on 6 August, Grotzinger says that he is "overjoyed" with the landing site, a level plain between the rim of Gale Crater and Aeolis Mons, the imposing mountain at the crater's centre. But he is sobered by the demands of managing such a complicated machine and the 400-strong science team behind it. "I feel the burden of two-and-a-half billion dollars," says Grotzinger, a geologist at the California Institute of Technology in Pasadena. "I feel the burden of the future of Mars exploration."

Grotzinger's job is to get the maximum scientific pay-off from that investment during the two years or more that the rover is expected to remain active. His day will begin at 12:42 a.m. with a science discussion with the principal investigators for the rover's ten instruments. But first Grotzinger needs breakfast and some coffee. One cup will do: the excitement of new data, which are expected in a downlink in two hours' time, will be enough to keep him alert.

Grotzinger was a regular at Conrad's in 2004, before and after his working days on the rover Opportunity, which landed that year along with Spirit, its twin, comprising the Mars Exploration Rover mission. Because the rovers were positioned on opposite sides of Mars, one team would be having breakfast while the other would be eating dinner. "The waitresses were always confused," he recalls. This time there is only one rover, but still no standard working day. Adapting to 'Mars time' requires starting each Earth day 40 minutes later than the last to match Martian daylight, inducing a state of perpetual jet lag.

Grotzinger is already marshalling different opinions on where the rover should go for its inaugural drive, expected to take place in the next week. The first images from the rover's navigation cameras suggested one possibility when they revealed

↵ NATURE.COM
For more on Curiosity, see:
nature.com/curiosity

The first high-resolution colour panorama from Curiosity's mast camera reveals the distant rim of Gale Crater.

whitish linear features in a shallow trench scoured out by the blast of the landing system's retrorockets. The lines could be fractures filled with water-altered minerals — key targets for exploration in the mission's goal of assessing whether ancient Mars could have supported life. "We want to find out if we've got a bird in the bush right there," Grotzinger says.

Several hundred metres away is a second option for exploration: a five-metre-high slope, or scarp, that separates the dusty desert pavement on which the rover landed from the more hummocky terrain that lies beyond. The scarp marks the edge of materials deposited at the foot of an alluvial fan, a sedimentary feature left behind from floods that once spilled down from the rim of Gale Crater. The materials seem to be cemented or consolidated in some way, and layers exposed on the scarp might reveal the comings and goings of an ancient lake.

Away in the distance is the rover's ultimate destination: the 5.5-kilometre-high Aeolis Mons, dubbed Mount Sharp by the science team, which casts a long shadow over the rover each morning. Data from orbiting spacecraft suggest that deposits on the mountain's lower shoulders were formed in water — and for now at least, these deposits represent the team's best chance of working out whether microbes could have lived on Mars billions of years ago.

Managing the tension between short-term and long-term goals will be a constant challenge for Grotzinger, says Raymond Arvidson, deputy principal investigator for the Spirit and Opportunity rovers and a planetary scientist at Washington University in St Louis, Missouri. "Do we stop and smell the roses or do we button up and continue to the next outcrop?"

Colleagues say that, in his geological field studies on Earth, Grotzinger balances both tendencies, displaying an exuberant fitness

that carries him across wide terrain even as he keeps an eye on the minutiae of mineralogy and texture. He scrambles like a mountain goat, says Ralph Milliken, a former postdoctoral student of Grotzinger's and now a planetary scientist at Brown University in Providence, Rhode Island. Yet Andrew Knoll, a palaeontologist at Harvard University in Cambridge, Massachusetts, recalls a field trip with Grotzinger to Siberia in the early 1990s, during which the geologist discovered small pieces of volcanic breccia rock underfoot that had gone unnoticed by Knoll, no slouch of a

Geologist John Grotzinger is plotting a course for Curiosity.

field geologist himself. By dating those rocks, the researchers were able to pin down the timing of the Cambrian boundary better than ever before, and show that the Cambrian explosion of animal life happened incredibly quickly (S. A. Bowring *et al. Science* **261**, 1293–1298; 1993).

Having an eye for detail and a gift for stratigraphic field work will certainly help Grotzinger to decode Mount Sharp and the millions of years of history that its layers represent. But he faces sociological challenges as well as scientific

ones. Unlike Steven Squyres, the charismatic planetary geologist from Cornell University in Ithaca, New York, who was principal investigator for Spirit and Opportunity, Grotzinger does not have final authority over plans for the rover. Instead, as project scientist, he heads an executive committee that is made up of a representative from NASA headquarters and the principal investigators for the ten instruments. "He appreciates that he's the head of a committee, and not the Pope," says Knoll.

Arvidson says that Grotzinger is already aware of the job's limitations. "John's working the crowd," he says. "He really understands that in order to get buy-in, you have to establish trust."

One exercise that Grotzinger is proud of is a crowd-sourced geological map of Curiosity's landing ellipse that was put together before touchdown. Recognizing that little was known about the fine-grained detail of the location, he divided the ellipse into 151 quadrangles, measuring roughly 1.6 kilometres on each side, and asked for volunteers from the science team to turn orbiter data into geological maps of each sector. It was partly a team-building effort in the lead up to the landing — how better to bring opinionated geologists into the fold than by working on a patchwork quilt? But he also wanted to show the wider team that the landing terrain was full of intriguing detail that should not be overlooked in a fast trek to Mount Sharp.

"There is going to be something for everyone," says Grotzinger, slurping his coffee and smiling. "I quoted Bob Dylan when I was talking to these guys … 'A lot of people don't have much food on their table. But they got a lot of forks 'n' knives. And they gotta cut somethin'.' If you have an empty plate, people are going to go for each other. But if there's a wealth of riches, we can all share." ■

The university at Târgu-Mureş is at the centre of plagiarism accusations.

MISCONDUCT

Romanian scientists fight plagiarism

Researchers set up independent review panel after misconduct scandals hit government.

BY ALISON ABBOTT

Plagiarism scandals have shaken Romania this year. Former research minister Ioan Mang[1] and current Prime Minister Victor Ponta[2] have both been implicated, and leaders at a large medical university in Târgu-Mureş face similar accusations — although the government seems to have ignored those charges. Such scandals have convinced some Romanian scientists that they need to fight back against a culture of plagiarism that they see as ingrained in the university system.

This week, researchers are launching an online service called Integru, which will investigate and expose cases of plagiarism and other academic misconduct in Romania. Each case will be accompanied online by commentaries from international — and independent — reviewers selected for their expertise in the relevant field.

Integru's stated goal is to "help reform and restore confidence in the Romanian research and education system". Its editorial board, which comprises Romanian scientists working inside and outside the country, will remain anonymous to avoid personal attacks from those accused of misconduct.

Some Romanian academics facing allegations of plagiarism "have tried to dodge the bullet by questioning the motivation of the accusers", says one Integru board member. For example, both Mang and Ponta have belittled the charges against them by arguing that the whistle-blowers were politically motivated. Integru should be able to thwart that strategy, says the board member, "because suspicious papers will be posted alongside documentation of plagiarized and copy–pasted sections,

and commented on by credible named experts from credible academic institutes around the world — the misconduct will be hard to refute".

Scientists concerned about plagiarism had once pinned their hopes on Romania's National Ethics Council, which last year was given a strengthened mandate to investigate wrongdoing under a law aiming to raise standards in universities. The council published decisions on 15 cases, and reached conclusions on another three. But many lost faith in its fairness when Ponta's government dismissed all the council's members on 8 June, just as they were poised to vote on the plagiarism charge against Mang (who had already resigned his ministerial post). The council was reconstituted with new members appointed by the government, and its first act was to clear Ponta of his plagiarism charges in July. However, an ethics commission established by the University of Bucharest, which had awarded Ponta his PhD, contradicted that verdict, saying that the prime minister was guilty as charged[3].

On the day it was dissolved, the former research-ethics council submitted to the science ministry its report on charges of plagiarism against three leading figures at the University of Medicine and Pharmacy of Târgu-Mureş: rector Leonard Azamfirei; former rector Constantin Copotoiu, who is now president of the university senate; and senate member Klara Brânzaniuc, a former vice-rector for research. The ministry has not published the report and did not respond to *Nature*'s request for information about it.

Marius Echim, a physicist at the Belgian Institute for Space Aeronomy in Brussels and vice-president of the dissolved council, chaired the investigation. He says that it "identified clear plagiarism in all three cases and recommended that two of the three professors involved be dismissed from the university". No action has been taken, and the apparent suppression of the case is disturbing, he says. While Echim was on the council, he adds, it saw cases of extensive plagiarism across many different disciplines in Romania.

The Târgu-Mureş cases refer to papers in the university's English-language journal *Acta Medica Marisiensis*, which contain sections identical to other researchers' works. Stefan-Virgil Hobai, a biochemist who last year retired from the university, e-mailed university faculty members in December 2011 with evidence of

---

MORE ONLINE

the plagiarism, which *Nature* has seen. Since then, he has identified a total of 22 cases of apparent plagiarism in *Acta Medica Marisiensis* and 14 cases in another of the university's journals, the English-language *Romanian Review of Laboratory Medicine*. That represents 7% of the papers he examined in each journal.

Peter Hantz, a Romanian physicist of Hungarian ethnicity based at the Friedrich Miescher Institute in Basel, Switzerland, submitted the Copotoiu, Azamfirei and Brânzaniuc cases to the National Ethics Council in January. Azamfirei told *Nature* that he thought Hantz, a prominent campaigner for Hungarian-language teaching in Romania, may have been trying to discredit the ethnic Romanian university leaders. Târgu-Mureş has a large Hungarian population and a history of ethnic tensions.

Sanda-Maria Copotoiu, editor-in-chief of *Acta Medica Marisiensis* (and the wife of Constantin Copotoiu), responded to the allegations by publishing an 'expression of concern' in the latest issue of the journal, referring anonymously to Hobai as a "character who styled himself as a prominent researcher and voluntary justice maker", whose "vicious" allegations are "intoxicating the net". She told *Nature* that all three accused researchers had asked her to remove their names from the papers long before the plagiarism charges reached the ethics council, because they became aware of the papers' existence only after publication. Their names have not yet been withdrawn, but Constantin Copotoiu's paper has been removed from the journal's website.

Constantin Copotoiu and Brânzaniuc did not respond to *Nature*'s request for comment. Azamfirei told *Nature* that there is no plagiarism in the article with his name on it, and that the National Ethics Council was not the appropriate body to review his case because he was not rector at the time the article was published.

Integru asked three reviewers to take a closer look at the Copotoiu case, and all concluded that the article did contain sections of copied text. One of the reviewers, Mihai Ionac, a vascular surgeon at the Victor Babes University of Medicine and Pharmacy in Timisoara, Romania, says that too many researchers in Romania do not take plagiarism seriously, and may not understand why it is classed as misconduct. The Integru initiative is "extraordinary, and comes at the right time, with all the scandals and turmoil now going on", he says. "Just a few years ago, no-one was even talking about plagiarism — it was covered up, like a cat covers up its dirt." ■ **SEE EDITORIAL P.253**

1. Abbott, A. *Nature* **485,** 289 (2012).
2. Schiermeier, Q. *Nature* **486,** 305 (2012).
3. Schiermeier, Q. *Nature* http://dx.doi.org/10.1038/nature.2012.11047 (2012).

Workers at a biosafety-level-3 lab in Entebbe diagnosed recent cases of Ebola.

INFECTIOUS DISEASES

Ebola outbreak tests local surveillance

Ugandan laboratory helps to minimize number of cases.

BY EWEN CALLAWAY

Film-goers know the script: a frightening infectious disease emerges somewhere in the developing world. Masked health-care workers arrive, blood samples are rushed to an airfield and the grim diagnosis is finally made in a state-of-the art laboratory on a different continent. But the response to the Ebola infections that have killed at least 16 people in Uganda since early July departed from the script — a laboratory in Entebbe, near Uganda's capital Kampala, has passed its first major test after taking a leading role in confirming and tracking the infection.

The lab, run by the US Centers for Disease Control and Prevention (CDC) and based at the Uganda Virus Research Institute (UVRI), tests tissue samples from patients with symptoms of the viral haemorrhagic fever. Its proximity to the outbreak (see 'Local hero') has helped to speed up the response, and the outbreak now seems to be under control. The episode highlights the advantages of having sophisticated surveillance labs in parts of the world where infections are likely to emerge, says Eddie Holmes, a geneticist who studies emerging infections at Pennsylvania State University in University Park. "The CDC making a commitment to this lab is a tremendous step forward."

The laboratory was set up after a 2007 outbreak of a newly discovered strain of Ebola called Bundibugyo, which killed 42 people. The area hit by the outbreak had no electricity and lacked the facilities to house a mobile containment lab, so scientists studying the disease set up shop at the UVRI. After the Bundibugyo outbreak subsided in 2008, the CDC invested in a permanent laboratory there that could diagnose haemorrhagic fevers such as Ebola and Marburg. Its biosafety-level-3 laboratory is safe enough to diagnose these viruses from tissue samples, but not to grow virus-infected cells for more sophisticated analyses, such as genome sequencing to study virus evolution. For that, samples are still sent to the CDC's headquarters in Atlanta, Georgia.

When the Entebbe lab isn't in outbreak mode — samples still arrive daily from the country's midwestern Kibale district — it can provide negative diagnoses for diseases with symptoms similar to those of viral haemorrhagic fever, such as bloody diarrhoea caused by shigellosis.

Uganda has had three Ebola outbreaks since 2000, so the lab should also help to catch future outbreaks early. "The earlier you catch an epidemic happening, the better you can stop its spread," explains Joseph Fair, vice-president of Metabiota, a company based in San ▶

the plagiarism, which *Nature* has seen. Since then, he has identified a total of 22 cases of apparent plagiarism in *Acta Medica Marisiensis* and 14 cases in another of the university's journals, the English-language *Romanian Review of Laboratory Medicine*. That represents 7% of the papers he examined in each journal.

Peter Hantz, a Romanian physicist of Hungarian ethnicity based at the Friedrich Miescher Institute in Basel, Switzerland, submitted the Copotoiu, Azamfirei and Brânzaniuc cases to the National Ethics Council in January. Azamfirei told *Nature* that he thought Hantz, a prominent campaigner for Hungarian-language teaching in Romania, may have been trying to discredit the ethnic Romanian university leaders. Târgu-Mureş has a large Hungarian population and a history of ethnic tensions.

Sanda-Maria Copotoiu, editor-in-chief of *Acta Medica Marisiensis* (and the wife of Constantin Copotoiu), responded to the allegations by publishing an 'expression of concern' in the latest issue of the journal, referring anonymously to Hobai as a "character who styled himself as a prominent researcher and voluntary justice maker", whose "vicious" allegations are "intoxicating the net". She told *Nature* that all three accused researchers had asked her to remove their names from the papers long before the plagiarism charges reached the ethics council, because they became aware of the papers' existence only after publication. Their names have not yet been withdrawn, but Constantin Copotoiu's paper has been removed from the journal's website.

Constantin Copotoiu and Brânzaniuc did not respond to *Nature*'s request for comment. Azamfirei told *Nature* that there is no plagiarism in the article with his name on it, and that the National Ethics Council was not the appropriate body to review his case because he was not rector at the time the article was published.

Integru asked three reviewers to take a closer look at the Copotoiu case, and all concluded that the article did contain sections of copied text. One of the reviewers, Mihai Ionac, a vascular surgeon at the Victor Babes University of Medicine and Pharmacy in Timisoara, Romania, says that too many researchers in Romania do not take plagiarism seriously, and may not understand why it is classed as misconduct. The Integru initiative is "extraordinary, and comes at the right time, with all the scandals and turmoil now going on", he says. "Just a few years ago, no-one was even talking about plagiarism — it was covered up, like a cat covers up its dirt." ■ **SEE EDITORIAL P.253**

1. Abbott, A. *Nature* **485,** 289 (2012).
2. Schiermeier, Q. *Nature* **486,** 305 (2012).
3. Schiermeier, Q. *Nature* http://dx.doi.org/10.1038/nature.2012.11047 (2012).

REUTERS/E. ECHWALU

Workers at a biosafety-level-3 lab in Entebbe diagnosed recent cases of Ebola.

Ebola outbreak tests local surveillance

Ugandan laboratory helps to minimize number of cases.

BY EWEN CALLAWAY

Film-goers know the script: a frightening infectious disease emerges somewhere in the developing world. Masked health-care workers arrive, blood samples are rushed to an airfield and the grim diagnosis is finally made in a state-of-the art laboratory on a different continent. But the response to the Ebola infections that have killed at least 16 people in Uganda since early July departed from the script — a laboratory in Entebbe, near Uganda's capital Kampala, has passed its first major test after taking a leading role in confirming and tracking the infection.

The lab, run by the US Centers for Disease Control and Prevention (CDC) and based at the Uganda Virus Research Institute (UVRI), tests tissue samples from patients with symptoms of the viral haemorrhagic fever. Its proximity to the outbreak (see 'Local hero') has helped to speed up the response, and the outbreak now seems to be under control. The episode highlights the advantages of having sophisticated surveillance labs in parts of the world where infections are likely to emerge, says Eddie Holmes, a geneticist who studies emerging infections at Pennsylvania State University in University Park. "The CDC making a commitment to this lab is a tremendous step forward."

The laboratory was set up after a 2007 outbreak of a newly discovered strain of Ebola called Bundibugyo, which killed 42 people. The area hit by the outbreak had no electricity and lacked the facilities to house a mobile containment lab, so scientists studying the disease set up shop at the UVRI. After the Bundibugyo outbreak subsided in 2008, the CDC invested in a permanent laboratory there that could diagnose haemorrhagic fevers such as Ebola and Marburg. Its biosafety-level-3 laboratory is safe enough to diagnose these viruses from tissue samples, but not to grow virus-infected cells for more sophisticated analyses, such as genome sequencing to study virus evolution. For that, samples are still sent to the CDC's headquarters in Atlanta, Georgia.

When the Entebbe lab isn't in outbreak mode — samples still arrive daily from the country's midwestern Kibale district — it can provide negative diagnoses for diseases with symptoms similar to those of viral haemorrhagic fever, such as bloody diarrhoea caused by shigellosis.

Uganda has had three Ebola outbreaks since 2000, so the lab should also help to catch future outbreaks early. "The earlier you catch an epidemic happening, the better you can stop its spread," explains Joseph Fair, vice-president of Metabiota, a company based in San ▶

Francisco, California, that conducts research on emerging diseases for the US government.

In May 2011, for example, a 12-year-old girl in Uganda was admitted to hospital with a fever and bleeding from nearly every orifice. She died within hours, and doctors suspected she had been infected with a virus such as Ebola or Marburg. In the past, a conclusive diagnosis would have meant chartering a plane to airlift samples to high-containment labs across Africa or in the United States, which can take weeks. Instead, a blood sample from the girl took just three days to arrive at the Entebbe lab, which confirmed Ebola. Health-care workers reached the girl's village a day later and helped to nip a would-be outbreak in the bud by tracking anyone who might have been exposed (T. Shoemaker *et al. Emerg. Infect. Dis.* http://doi.org/h53; 2012).

Without firm diagnoses, workers tracking outbreaks have to keep tabs on thousands of suspected or potential cases. But having the

LOCAL HERO
The Uganda Virus Research Institute was well placed to offer rapid diagnoses for suspected Ebola cases in midwestern Kibale.

AFRICA

Ebola outbreak zone

UGANDA

Entebbe • Kampala

Lake Victoria

Diagnostic lab

100 km

CDC lab so close to the outbreak means that "we can go back and forth to the site of the epidemic every day", says Pierre Rollin, a virologist at the CDC's Special Pathogens Branch in

Atlanta, who is currently in Uganda. "Any samples collected this morning in Kibale will arrive at the lab and be tested tomorrow." In the latest outbreak, rapid diagnosis has meant that health-care workers had to track mere dozens of suspected cases.

"Other countries are asking us if we can develop the same kind of lab for them, but we don't have the resources to do it," says Rollin. Instead, the CDC and other Western funders support programmes to build and run laboratories in the developing world and train their staff, so that countries can monitor and respond to outbreaks with their own resources. Gabon, for example, operates a maximum biosafety-level-4 laboratory.

"You want to build capacity in the country," says Stephen Morse, an epidemiologist at Columbia University in New York, "and not wait until they ask for help to send people parachuting in." ∎

CHEMISTRY

Amino acid provides shortcut to drugs

Organocatalyst halves synthesis of prostaglandin family.

BY KATHARINE SANDERSON

An iconic but convoluted pharmaceutical synthesis received a major overhaul this week. The approach, described online in *Nature*, gives researchers easier access to a family of molecules that includes a blockbuster drug.

The efficient new route to prostaglandins, a class of hormone-like chemicals that influence processes including blood circulation and inflammation, offers a way to churn out many different versions of the molecules. It also marks a victory for one of the hottest fields in organic chemistry: organocatalysis, which speeds up chemical reactions with the help of carbon-based molecules, potentially avoiding some of the drawbacks of conventional metal catalysts.

Prostaglandin-like drugs are tempting targets for medicinal chemists; one example, Pfizer's glaucoma drug latanoprost, achieved annual sales of around US$1.75 billion before its patent expired in 2011. But it takes an arduous 20 chemical steps to make latanoprost using the original strategy developed in 1969 by Elias Corey at Harvard University in Cambridge, Massachusetts. The synthesis helped Corey to win the Nobel Prize in Chemistry in 1990.
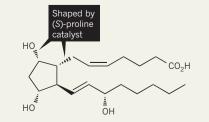
Synthetic chemist Varinder Aggarwal at the University of Bristol, UK, realized that he could

radically simplify the process by building the heart of a prostaglandin molecule in a different way. He applied a workhorse of organic chemistry, the aldol reaction, to build key carbon–carbon bonds in a prostaglandin's five-atom ring, but used the amino acid (*S*)-proline as an organocatalyst to make sure that the end product had precisely the right shape (G. Coulthard *et al. Nature* http://dx.doi.org/10.1038/nature11411; 2012). Corey's approach takes nine steps to make a different intermediate with the same crucial ring.

The molecule produced in Aggarwal's reaction serves as a foundation stone for the whole

ON TARGET
Prostaglandin PGF$_{2\alpha}$ (below) can now be made in a seven-step synthesis that uses a simple catalyst to orient key bonds after just two steps, producing an intermediate that could generate many other prostaglandins.

Shaped by (*S*)-proline catalyst

HO

HO

CO$_2$H

OH

family of prostaglandins, and his method can produce grams of this building block at a time — a viable scale for drug-development work. Aggarwal and his colleagues used it to make the most complicated prostaglandin, PGF$_{2\alpha}$, which is very similar to latanoprost, in just seven steps (see 'On target'). It's "by far the single best approach to the prostaglandins", says Ben List, an expert in organocatalysis at the Max Planck Institute for Coal Research in Mülheim an der Ruhr, Germany. As well as being shorter than Corey's version, it is very specific at producing only left-handed or right-handed versions of the molecules — a crucial factor in making an effective drug.

Whether the synthesis helps to deliver new drugs remains to be seen, says biomedical researcher Colin Funk at Queen's University in Kingston, Ontario, Canada. "Prostaglandin-like drugs are still relevant, but the big thing has been to try and limit their side effects," he says. That demands a better biological understanding, rather than a new synthetic route, he notes.

The method might prove useful in full-scale drug production, not only because of its simplicity but because it avoids expensive and toxic metal catalysts. "From an industrial perspective Aggarwal's process is promising, and a remarkable lead," says process chemist Rob Singer, who works at Pfizer Worldwide Research and Development in Groton, Connecticut. The synthesis would still need further development before it could be used at a production scale, he adds.

List says that most pharmaceutical companies are studying organocatalysts, but so far have tended to use them on only a small scale. As the chemical structures of drugs get more complicated, he hopes that organocatalysts will gain more prominence, and Aggarwal's work should help to speed the change, he says. "It reflects the way that organocatalysis fundamentally changes the way we do catalysis," he says. ∎

AIDS

Generic HIV drugs will widen US treatment net

Upcoming patent expiries stand to make medicines cheaper — although less convenient.

BY AMY MAXMEN

HIV remains a devastating scourge worldwide — and an expensive one. In the United States, where HIV care costs the most, the annual price of treatment averages almost US$20,000 per patient, and the federal government foots most of the bill. This year alone, the United States will spend nearly $15 billion on HIV care and medication, yet funding shortfalls mean that more than 2,000 patients remain on a waiting list for antiretrovirals, and many others are not receiving treatment for co-infections such as hepatitis C.

In the next few years, that economic burden looks set to ease as US patents on the 'superstar' HIV drugs begin to expire (see 'Time's up'). Generic versions of brand-name antiretrovirals, whose use is permitted in the developing world, could reduce the cost per patient and allow the US budget to stretch further. "AIDS, as we currently know it, may become much more affordable," says John Bartlett, an infectious-disease researcher at Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland.

Earlier this year, the HIV/AIDS working group at the Department of Health and Human Services asked Rochelle Walensky — an infectious-disease physician at Massachusetts General Hospital in Boston who is known for her cost-effectiveness analyses — for an opinion on the impact of a switch from Atripla, a widely used 3-in-1 HIV pill, to a combination of generic and brand-name drugs might be. Atripla is the first of four first-line drug combinations recommended by US guidelines that will be affected by patent expiries. A key component of the pill will lose protection next year.

At last month's International AIDS Conference in Washington DC, Walensky announced that the switch would save the United States a projected $920 million in the first year alone. She estimates that Atripla — a co-formulated pill from Gilead Sciences in Foster City, California, that consists of brand-name emtricitabine, efavirenz and tenofovir — costs about 40% more than an alternative three-pill regimen. This would include: generic lamivudine, which can be used as a substitute for emtricitabine; a generic version of efavirenz that is expected to hit the market once the brand goes off-patent in 2013; and brand-name tenofovir.

In her presentation, Walensky estimated that the replacement could be slightly less effective than Atripla, not least because patients are more likely to skip doses when they have to swallow several pills. But the difference is small, according to Walensky's model: staying with Atripla rather than switching to the three-pill regimen would add 4.4 months to a patient's life, but $42,500 to their bill.

Even in the United States, where discussions of cost-effectiveness in health care are politically charged, this cost-benefit ratio may be too high for any health-care innovation. "I'm a doctor and I would never want to put a patient on a lesser drug — but we do have to think about trade-offs," Walensky says. In this case, she says, the trade-off could be between a slightly less-effective treatment for the individual but greater care for people with HIV overall.

Patient groups are not immune to this reasoning. "Although there's a convenience factor to having one pill once a day, if the trade-off is many people having no access to medication, then it's better to split up the pill or use the threat of splitting it to get Gilead to reduce their prices," says Michael Weinstein, president of the AIDS Healthcare Foundation, an advocacy organization based in Los Angeles, California. Gilead would not comment directly on pricing, but a spokeswoman told *Nature*, "We believe there is a lot of data supporting the value of single-tablet regimens".

Analogous questions will arise as several other first-line medicines lose patent protection over the next five years. If generic versions are embraced in the United States, drug companies will feel a squeeze that they say could inhibit the development of new HIV drugs. But some say a widespread switch is unlikely to happen, noting that newer patented drugs are on the way, including Gilead's Quad, a combination therapy slated for approval by the US Food and Drug Administration on 27 August. This 4-in-1 pill contains a novel integrase inhibitor, and outperformed Atripla in clinical trials. If the bulk of patients switch directly from Atripla to Quad, Gilead will retain its hold on the US market.

Bartlett argues, however, that the semi-generic substitute may become popular among the thousands of people who currently control HIV with Atripla, with Quad reserved for those in need of an alternative. With so many drugs going off-patent, he predicts, "In ten years this will be a disease treated for $200 per year, or less". ∎

TIME'S UP
Patents on several key HIV drugs are set to expire over the next five years

Drug	Type	Manufacturer	Expiry
3TC (lamivudine)	Nucleoside reverse transcriptase inhibitor	GlaxoSmithKline	2010
Abacavir	Nucleoside reverse transcriptase inhibitor	GlaxoSmithKline	2012
Efavirenz*	Non-nucleoside reverse transcriptase inhibitor	Bristol-Myers Squibb	2013
Delavirdine	Non-nucleoside reverse transcriptase inhibitor	Pfizer	2013
Darunavir	Protease inhibitor	Tibotec	2015
FTC (emtricitabine)*	Nucleoside reverse transcriptase inhibitor	Gilead	2015
Tipranavir	Protease inhibitor	Boehringer Ingelheim	2015
Ritonavir	Protease inhibitor	Abbott	2016
Tenofovir*	Nucleotide reverse transcriptase inhibitor	Gilead	2017

*Component of Atripla

TASTE FOR ADVENTURE

Scientists have travelled the ancient Silk Road to understand how genes shape people's love for foods.

BY ALISON ABBOTT

It wasn't quite the lowest point of the expedition. But Paolo Gasparini was decidedly uncomfortable at a banquet in the small town of Alga, Kazakhstan, when he was given the honour of dividing a boiled sheep's head among the gathering. He could sense the silent alarm of his young colleagues, fearing that soon they would be politely chewing on an eye or an ear. Fighting nausea, Gasparini, a medical geneticist at the University of Trieste, Italy, sympathetically sliced each of them just a taste of skin, passing the traditionally valued sense organs to the village elders.

The sheep's-head ritual did, however, neatly reflect the mission of the trip. The Italian scientists were gathering DNA samples from members of the many different ethnic groups that live along the Silk Road — a 2,000-year-old

trading route between China and Europe — to map the genetic roots of food preference and of the senses that contribute to it. The team set out from Trieste in July 2010 and collected its final samples this June in northern Kyrgyzstan. It had amassed about 1,100 DNA samples and thousands of sensory-test results from a dozen or so distinct populations, ranging from Caucasian to Turkish and Mongolian.

FULL FLAVOUR

Over the past two decades, researchers have identified receptors on the tongue for sweet, umami ('savoury'), bitter, sour and salty. They have realized that the sensation of flavour is created when signals from these receptors are integrated in the brain with food-related signals from the hundreds of odorant receptors

in the nose and nasal passages. But understanding why people prefer certain foods is still a major scientific puzzle.

Subtle variations in the genes that encode the taste and smell receptors are likely to have a big influence. But many genes involved in touch, hearing and sight — those that respond to the creamy feel of fatty food, the sound of crunching nuts, the sight of a crimson tomato — also have a role. Travelling the Silk Road gave Gasparini and his team a unique opportunity to identify genetic influences on food preference by studying isolated populations, in which such variants are relatively easy to find. "The expedition approaches an almost unexplored

SILK-ROAD SCIENCE

Over three expeditions in three summers, researchers carried out tests on ethnic groups along the Silk Road, the 2,000-year-old trading route between China and Europe. The aim is to map the genetic roots of food preference.

Trieste
Expedition start

Collecting saliva for DNA

Testing hearing

Indicating food preferences

Testing smell

dimension of what makes people different," says Yurii Aulchenko, a statistical geneticist from the Russian Academy of Sciences Institute of Cytology and Genetics in Novosibirsk.

Researchers in evolutionary and population genetics are watching with interest, as is the food industry. Humans evolved in environments where food was scarce and poisonous plants a constant threat. As a result, calorie-dense fatty foods can be hard to resist, whereas the bitter tastes that are common in toxic plants can make some nutritious foods, such as broccoli, unpalatable. In the modern, food-rich world, that genetic legacy helps to push people towards diets that contribute to obesity and ill-health. Understanding the roots of food preferences might one day help to change people's behaviour. "Eventually, industry might be able to develop foods that are enjoyable without being bad for you," says Jim Kaput, head of clinical translational research at the Nestlé Institute of Health Sciences in Lausanne, Switzerland, which collaborates with the Nestlé Research Center in Lausanne on studies linking genomics and food preferences. "But right now we understand almost nothing — we need a lot more basic research," says Kaput.

GENETIC GASTRONOMY

The Silk Road expedition — officially called Marco Polo, after the Venetian explorer who travelled the route in the thirteenth century — was born five years ago, when Gasparini met Enrico Balli, director of the Medialab of the International School for Advanced Studies in Trieste. The two hit it off immediately. Balli,

a straight-talking physicist turned science-communicator, was itching to film exotic locations. Gasparini, tall and genial, had spent most of his scientific life studying the genetics of isolated communities in dead-end valleys of the Italian Alps. He had become interested in the genetics of food preferences, was consulting with food companies, and had been dreaming vaguely of accessing remote, isolated populations in more grandiose mountain chains in Asia or South America, where scientists have rarely ventured.

Such populations were often founded by just a few people who went on to marry within the group, so their descendants tend to be genetically homogeneous. This makes it easier to work out if a gene variant is associated with a particular characteristic, such as a taste preference, than it is in a large, genetically diverse group. Comparing different isolated populations can be even more valuable. If a variant is associated with the same trait in genetically different groups, that may be a sign that it is important across almost all populations.

The Silk Road offers a potential paradise for such genetic exploration. The route traverses massive mountain ranges such as the Pamir and the Tian Shan in central Asia and passes through pockets of the nomadic tribes who originally populated the region, as well as ethnically diverse groups descended from traders who settled en route, often near the roadside inns called caravanserais. These populations did not tend to share their genes, but they did share recipes. Cuisines are remarkably similar along much of the Silk Road — variations on

tandoor breads, noodles with vegetable or mutton sauces, and dried or fresh fruit. This means that differences in food preferences between groups are likely to be down to variations in genes rather than in dietary cultures, making them even more appealing to the geneticists.

But before Gasparini and Balli could study these groups, they had to reach them. They originally conceived the expedition as a do-it-yourself venture, in which they would simply drive themselves from country to country. Fortunately, they consulted Lilia Smelkova — co-founder of Terra Madre, a network of farmers in developing countries established by the non-profit campaign group Slow Food, which is based near Turin, Italy. Smelkova is an expert in the former Soviet republics that Gasparini and Balli planned to visit. "I thought they were quite crazy," she says, knowing the civil-war-torn, multilingual, sometimes despotic territories that the scientists planned to traverse and the closed communities that they wanted to study. "But I loved their ideas, so I agreed to help — by joining them." Smelkova used her contacts to help the researchers to win the confidence, and the cooperation, of villages on the route.

TASTE TEST

The team, consisting of ten or so people, split the expedition across three summers (see 'Silk-Road science'). The first year, they set off for Georgia armed with just a few clothes and some compact equipment, and fell into a gruelling schedule. Each day, they drove — often for many hours — to a community that was to be sampled and tested. A local Terra

Travelling between villages

Testing taste

KAZAKHSTAN

KYRGYZSTAN

TAJIKISTAN

CHINA

Ancient Silk Road route

Shanghai
Expedition finish

● Sampling point

Madre helper explained the tests, and volunteers were asked to donate DNA by spitting into plastic tubes. Then the scientists tested volunteers' abilities to hear different sound frequencies, to taste salty or bitter compounds soaked into scraps of filter paper, to distinguish different shades of colour and to identify 12 different smells impregnated in marker-pen-like sticks. The team also showed the volunteers pictures of 80 different food types and asked them to indicate how much they liked or disliked each one. At night, the scientists would feed the results into their computers before collapsing into bed.

The expedition's success depended entirely on getting sample tubes containing saliva back to Italy, but this proved nerve-rackingly difficult. When they arrived home after the first leg in September 2010, the researchers found that Uzbek customs had stopped a precious shipment of 350 tubes gathered in Azerbaijan, Turkmenistan and Uzbekistan. Despite exhaustion, Balli says, "I flew back to Tashkent, and brought them straight back in a suitcase."

Systematic analyses of the Silk Road data will keep the scientists busy for years. The team is carrying out a detailed analysis of small variants known as single-nucleotide polymorphisms (SNPs) in all the DNA samples, and will fully sequence up to 50 samples.

The scientists have already identified eight variants in known genes, including one for an ion channel involved in sensing spicy-hotness, which are associated with a taste for particular foods. And they have found that variants of the gene for the TAS1R2 protein, part of a sweetness receptor, are associated with a strong liking for vodka and white wine (N. Pirastu *et al. J. Food Sci.*; in the press). The results fit with reports linking preferences for alcohol and sweetness in mice (S. M. Brasser *et al. Physiol. Genomics* **41**, 232–243; 2010), and hint that both substances are perceived by the same receptor.

There may be bigger scientific stories hiding in the data. Gasparini says that the team is seeing an emerging association in Tajikistani populations between an olfactory receptor gene and both sensitivity to bitter tastes and a tendency to mistake smells. If the finding holds up, it will be the first demonstrated genetic link between smell and taste perception, and it could help to explain how signals from different senses combine to sculpt individual food preferences. The finding may also offer clues to how evolution might have shaped the senses. "Who knows?" Gasparini muses. "Might evolution have compensated those who can't distinguish smells efficiently by selecting genes that make them more sensitive to bitter tastes, so they have a better chance of recognizing poisons?"

DIET AND DIVERSITY

The genetic data will serve another purpose by filling a gap in the international 1000 Genomes Project, which is cataloguing genetic variants in different populations across the globe, but until now had no representatives from central Asia. Population geneticists curiou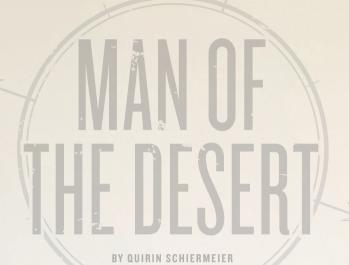s about how different peoples evolved will find their own stories in the Silk Road DNA, says Aulchenko. "We know a lot about the big migrations — out of Africa, for example — but we don't know how trade shaped genetic variation."

The data were hard won. At the trip's lowest point, on the drive through Tajikistan from Qal'aikhum to Khorugh in the Pamir Mountains, a road closure forced the jeep convoy to detour for 16 tense, sweltering hours along the Panj River, which forms part of the border with Afghanistan. The few metres between the road and the river were laid with landmines.

Still, Gasparini and Balli have caught the fieldwork bug, and are planning another expedition for next year. At the end of the nineteenth century, members of a small, isolated Italian population near Trieste emigrated in droves to South America — where they settled in different valleys in the Andes. Gasparini wants to compare these people's taste preferences and the acuity of their senses with those of their relatives in Italy. Unlike the Silk Road populations, the groups have the same genetic backgrounds but very different diets. Gasparini hopes that this will create a natural laboratory for exploring the interaction of diet and other environmental factors with interesting gene variants identified in the Silk Road samples.

Any discomforts of expedition science are overwhelmingly compensated for by the gains, says Gasparini. And if nothing else, the Marco Polo expedition taught him that boiled sheep's head is not his dish of choice. ∎

Alison Abbott *is* Nature's *senior European correspondent.*

MAN OF
THE DESERT

BY QUIRIN SCHIERMEIER

Stefan Kröpelin stands near
one of Chad's Ounianga lakes,
an area that he helped to get
listed as a World Heritage site.

Stefan Kröpelin has carved out a career where few dare to tread — in the heart of the Sahara.

Stefan Kröpelin's first research trip to the eastern Sahara was nearly his last. As a graduate student in January 1982, he travelled alone to Gilf Kebir — a remote plateau in the southwest corner of Egypt — to study sediment deposition in the steep canyons. Things didn't start well. On the first night, he found that his thin clothes and sleeping bag were no match for the freezing conditions. During the day, fierce sandstorms forced him to dig a pit for shelter and keep his eyes and mouth shut for most of the time. After a week, the driver who was scheduled to pick him up was nowhere in sight. Three days later, and running out of water, Kröpelin was just about to embark on a desperate 150-kilometre march to the nearest oasis when the jeep finally arrived.

Kröpelin, now a geologist and climate researcher at the University of Cologne in Germany, not only survived the ordeal but went on to become one of the most devoted Sahara explorers of our time. In dozens of expeditions to some of the farthest corners of the desert (see 'A career in the sand'), he has endured week-long sandstorms, suffered bouts of serious disease, including schistosomiasis, and faced heavily armed groups that roam the eastern Sahara.

But those decades of difficult field work have paid off for Kröpelin, who has made seminal discoveries about the climatic history of the Sahara that are challenging assumptions about the tipping points the world may face in a warmer future. At the same time, Kröpelin has worked to document the Sahara's cultural history and to preserve its heritage by lobbying to protect important scientific and cultural sites. "I've never ceased to be excited about the chance to investigate some of the least known parts of the desert," he says.

Siddiq Abd Algadir, president of the Sudanese Geologists' Union in Khartoum and a fellow student with Kröpelin in the 1980s, says that the German researcher has added immeasurably to Saharan science. "Much of what we now know about the geology, the environments and even the people in some of the most remote parts of the Sahara, we really owe to him and the expeditions he has led."

With his wiry build and boundless energy, Kröpelin seems much younger than his 60 years. The desks in his office are covered with parchment-coloured maps featuring vast expanses of desert. And his bookshelves sag with hundreds of volumes covering Africa's natural and cultural history, from early human evolution and migration to the trans-Saharan trade routes that carried gold and salt between West Africa and the Mediterranean.

In the 1960s, Kröpelin's sense of adventure was sparked by reading the exploits of nineteenth-century explorers such as Heinrich Barth and Gustav Nachtigal, who were among the first Europeans to document the cultures and environment of the Sahara. Two decades later, in graduate school at the Technical University of Berlin, that early curiosity about the area became a scholarly passion as he studied its physical geography and geology.

Similar in size to Australia or the United States, the Sahara is one of the most inhospitable and least explored spots on the globe. Throughout his work there, Kröpelin has tried to amass detailed knowledge of the scientific and cultural facets of his study areas before turning out results. Researchers who have travelled with Kröpelin describe him as a circumspect, hard-driving leader, equally at home in geomorphology, archaeology, human psychology and — when necessary — motor mechanics.

MAP QUEST

Over the past few years, Kröpelin's work has helped to reveal how the Sahara transformed from a savannah more than 5,000 years ago to the desert it is today. Researchers had previously thought that the transition happened abruptly — within little more than a century — when a cyclical shift in Earth's orbit reduced the amount of sunlight in the tropics and weakened the African monsoon. This idea was championed by Peter deMenocal of Columbia University in New York, who studied a core of deep-sea sediments from the tropical Atlantic ocean and saw a rapid rise in the amount of dust blowing off the continent at the time of climatic transition[1]. The sediment record agreed with models suggesting that a number of sensitive components of the planet — such as monsoon patterns, Amazon and Saharan vegetation, polar ice sheets and even the Atlantic Ocean circulation — will, under certain circumstances, flip from one stable state to another[2].

But Kröpelin wasn't convinced. The concept of an abrupt climate switch didn't mesh with his previous research on ancient settlements in the eastern Sahara[3]. "There is evidence from thousands of archaeological sites throughout the Sahara that prehistoric human settlements weren't abandoned within a few decades or so," he says. He was also piqued that deMenocal reached his conclusion without ever setting foot in the desert, and used a single marine record to make generalizations about the entire Sahara. "The idea of catastrophically fast climate change is untenable — it can only come from someone who doesn't know the Sahara," says Kröpelin.

From his past reconnaissance work, Kröpelin knew one place in the Sahara with a climate record potentially long enough to test his ideas — a series of permanent lakes in the remote Ounianga basin in northeastern Chad, a practically rainless and notoriously insecure region bordering Libya and Sudan. French military geographers had explored the lakes in the early twentieth century but the area had since been forgotten by most foreign researchers. The lakes are fed by ancient groundwater, and their depths are layered with sediments containing climatic indicators such as pollen. Kröpelin suspected that one of the largest lakes, Lake Yoa, held a sedimentary cache thick enough to reconstruct the long-term climatic history of the region.

Working in Ounianga has its challenges. On a reconnaissance mission in 1999, Kröpelin had to persuade a superstitious local population that drilling a hole in the lake bottom would not drain the water or release the jinns and demons they feared might dwell in the ground.
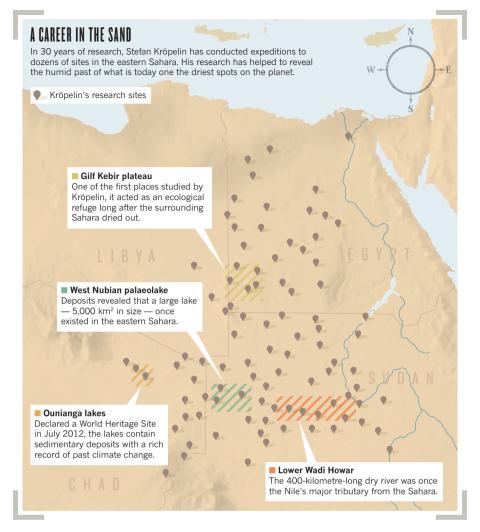
After extensive preparations, he returned to the remote lake in 2003 and 2004, and pulled up a continuous 8-metre-long core of sediment from its bottom — a climate record reaching back 6,000 years. Kröpelin's analysis revealed that it took some 3,000 years — from 5,600 to 2,700 BC — for the fully vegetated savannah there to transform into a barren desert[4].

THE GREAT DRYING

At first, some researchers questioned the dating of the core and whether it represented conditions across a broad region. But Kröpelin answered his critics[5], and his ideas eventually gained acceptance. Today, deMenocal maintains that the western Sahara dried out quickly, but he acknowledges that the eastern part around Lake Yoa took much longer. "Stefan knows the Sahara better than any other living scientist, and the climate archive he has uncovered, while not perfect, is uniquely well-dated," he says.

The debate has intrigued climate modellers interested in potential tipping points in the Earth system, in which small perturbations cause abrupt and major changes. Kröpelin's findings imply that the most recent savannah-to-desert transition does not fit this pattern — although he leaves open the possibility of earlier tipping points in the Saharan climate.

Climate researchers are therefore eager for Kröpelin to report his analysis of a 16-metre-long core from Lake Yoa that he drilled in 2010, which he is now readying for publication. "The published record is fascinating but rather

A CAREER IN THE SAND

In 30 years of research, Stefan Kröpelin has conducted expeditions to dozens of sites in the eastern Sahara. His research has helped to reveal the humid past of what is today one the driest spots on the planet.

📍 Kröpelin's research sites

Gilf Kebir plateau
One of the first places studied by Kröpelin, it acted as an ecological refuge long after the surrounding Sahara dried out.

West Nubian palaeolake
Deposits revealed that a large lake — 5,000 km² in size — once existed in the eastern Sahara.

Ounianga lakes
Declared a World Heritage Site in July 2012, the lakes contain sedimentary deposits with a rich record of past climate change.

Lower Wadi Howar
The 400-kilometre-long dry river was once the Nile's major tributary from the Sahara.

LIBYA   EGYPT   SUDAN   CHAD

SOURCE: S. KRÖPELIN

short," says Martin Claussen, a climate modeller at the Max Planck Institute for Meteorology in Hamburg, Germany, who is exploring ways to predict fast climate–vegetation transitions[6]. "The complete 12,000-year record should help us get a better sense of how rapidly the system might be capable of switching from one state to another and what the transition period may look like."

The results from Lake Yoa crown a long list of discoveries that Kröpelin has made in the region. In one of his earliest major finds, Kröpelin established that the dry valley known as Wadi Howar, which sits in an extremely arid part of northern Sudan, was once one of Africa's largest rivers and a tributary to the Nile[7]. This extinct river flowed from about 9,500–4,500 years ago and supported a rich savannah that was home to a host of animals, including antelopes, giraffes, zebras and elephants.

Kröpelin's extended trips through the Sahara, which typically last two months, have sometimes placed him in difficult situations. During one expedition to northern Sudan in 1995, he was chased for several days across the desert by a truck. When the pursuers caught him, they turned out to be

Sudanese forces who thought they were following outlaws smuggling weapons or people. The soldiers were nearly out of fuel from their chase, so Kröpelin gave them a share of the expedition's diesel supplies and they enjoyed a friendly meal together. The excavations carried out on that trip revealed how prehistoric humans settled there during the Sahara's humid phase and then disappeared as the climate dried out[3].

Another experience in 2005 was more troubling. One night, the sound of heavy gunfire terrified a team that Kröpelin had led to Erdi Ma, an uninhabited and largely unexplored tableland in the northwest corner of Chad. The next morning they learned that Sudanese military forces had massacred 20 Darfur rebels not far from the expedition's campsite.

Sometimes, nature is the biggest enemy. On the way to Lake Yoa in April 2010, Kröpelin and his team got caught in a dust storm fiercer than any he had seen in 30 years. For more than a week visibility dropped to two metres. Under non-stop bombardment, the researchers' bodies and vehicles became electrically charged and each touch caused painful electric shocks. Refilling the trucks from metal gas

canisters was Russian roulette, as large sparks shot around the filler necks. "Back home, it took weeks until I could open my car again without inhibition," Kröpelin says.

Through his expeditions, Kröpelin has developed relationships with many of the regional governments, and he has helped to advise officials in Egypt, Libya and Chad on the impact of mining, oil drilling and tourism. He has campaigned for and won protection for two environmentally and culturally unique regions that he explored — Wadi Howar and Gilf Kebir.

For the past 13 years, Kröpelin has also lobbied to protect the Ounianga lakes — an effort that involved hundreds of meetings with ministers in Chad and international agencies and officials. That work paid off in July, when the region was declared a world heritage site by UNESCO (the United Nations Educational, Scientific and Cultural Organization).

Researchers wonder what will happen when Kröpelin retires, leaving a major hole in Saharan studies. "He has taken risks few researchers would be happy to take any more," says deMenocal.

The risks in the region have grown lately because of rebel activity and recurring armed conflict in places including Darfur, Libya, Mali and Chad. But despite some anxious moments over the years, Kröpelin considers that the risk of getting robbed, kidnapped or killed is smaller in the Sahara than in sections of London or New York City.

"In the desert there is an air of fellowship with whoever crosses your way," he says. "Encounters with nomads, rebels and more shady folks can be scary, but I've always managed, somehow, to talk us out of even the most delicate incidents."

For now, Kröpelin shows no signs of slowing down. He plans to return to the Sahara in November for a celebration of the UNESCO decision and to do some geological field work. Next year, he will search for traces of prehistoric writing and trading, and will explore a mysterious crater in Jebel Uweinat, the tallest mountain in the eastern Sahara. Gone are the days of foolhardy solo trips, but not the deep love for the land. "Under the vast desert skies, your whole existence takes on a different meaning," he says. "I'd like to return for as long as I can." ■

Quirin Schiermeier *is a reporter with* Nature *in Munich.*

1. deMenocal, P. *et al. Quat. Sci. Rev.* **19**, 347–361 (2000).
2. Lenton, T. M. *et al. Proc. Natl Acad. Sci.* **105**, 1786–1793 (2008).
3. Kuper, R. & Kröpelin, S. *Science* **313**, 803–807 (2006).
4. Kröpelin, S. *et al. Science* **320**, 765–768 (2008).
5. Kröpelin S., Verschuren D. & Lézine A.-M. *Science* **322**, 1326c (2008).
6. Bathiany, S., Claussen, M. & Fraedrich, K. *Clim. Dyn.* **38**, 1775–1790 (2012).
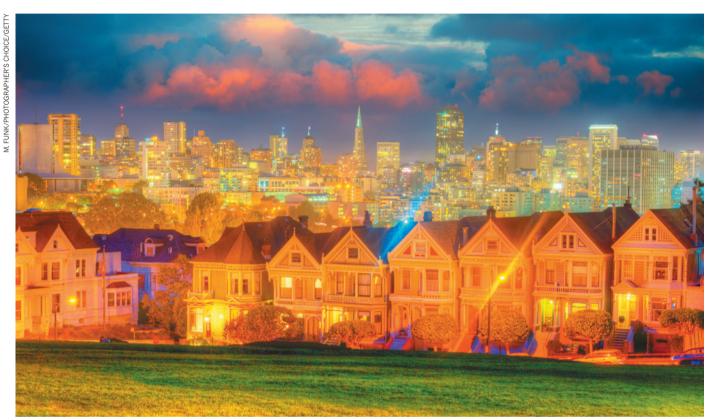7. Pachur, H.-J. & Kröpelin, S. *Science* **237**, 298–300 (1987).

COMMENT

Uses such as lighting, heating and cooling in buildings account for almost 40% of the energy consumed in the United States.
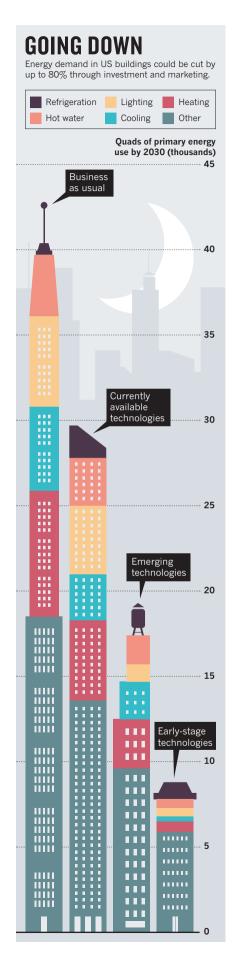
How to build a low-energy future

Advanced construction technologies promise huge energy savings, says **Philip Farese**. Investment is needed to bring them to market and to encourage their use.

In 2010, the United States used 98 quadrillion British thermal units (quads) of energy — about 20% of the world's total — on everything from transport to entertainment. A surprisingly large amount — 40 quads — went into buildings for uses such as space heating, cooling, lighting and computing. Assuming that business continues as usual, by 2030, the nation will require 43 quads of energy for buildings each year[1].

There is huge room for improvement without sacrificing an ounce of comfort. Much more should be invested in developing energy-saving technologies, in bringing them to market and in encouraging their use.

I, along with colleagues at the National Renewable Energy Laboratory (NREL) in Golden, Colorado, and at the US Department of Energy (DOE), have reviewed more than 300 publications and sought the advice of 65 experts to determine how much US energy demand for buildings might be cut. We find[2] that cost-effective technologies available now could reduce energy use in buildings by up to 30% by 2030, which is in line with others' findings[3].

Add in the affordable emerging technologies that are expected to be available in the next 5 years or so, and the potential reduction reaches 55%. Including all ▶

GOING DOWN

Energy demand in US buildings could be cut by up to 80% through investment and marketing.

Legend:
- Refrigeration
- Hot water
- Lighting
- Cooling
- Heating
- Other

Quads of primary energy use by 2030 (thousands)

- Business as usual
- Currently available technologies
- Emerging technologies
- Early-stage technologies

▶ technologies that are in the early stages of research and development (R&D) brings this cut in energy use to an astonishing 80% (see 'Going down').

Some researchers have pointed out that efficiency improvements can lead to an increase in demand[4]. Such 'rebound' effects are real, but are expected to soak up no more than 15% of energy savings[5]. Some people might heat their homes a bit more in the winter, but very few will wash more clothes or cook more meals. Other researchers argue[4] that the money saved through energy efficiency will be spent by consumers on energy-intensive luxuries, such as long-distance flights. But consumers are likely to spend in line with their usual budget. For every dollar spent in the United States, about 8% goes back to energy costs, including the energy used to manufacture, sell and transport goods — an indirect rebound effect.

Even accounting for these effects, energy savings should still exceed 20% from currently available technologies and 45% from emerging technologies[2]. The 80% reduction as a result of advanced technologies becomes 65% when rebound effects are included, meaning that the projected 43 quads of demand could be met using some 14 quads of energy. The national energy bill could be reduced by as much as US$275 billion a year[2] if just a fraction of that amount was invested strategically.

According to my analysis, energy-use intensity — the amount of energy used per unit area of building space per person — actually rose in the United States by 10% from 1980 to 2010. Because of this, some say that efficiency measures serve to increase energy use. But this rise results from an increase in demand for modern devices such as air-conditioning units and computers. Heating (of both water and space) and refrigeration have also seen a small rise in demand despite significant improvements in equipment efficiency — but by my calculation, their energy-use intensity has fallen by 43%. Efficiency will need to be pursued aggressively in all areas, using approaches that range from increasing R&D investment to maintaining appliance standards.

HEAT AND LIGHT

For centuries, most of the energy used to heat buildings came from burning fuels, which can provide no more than one unit of heat for each unit of energy. But many energy-efficiency technologies are in the pipeline, and a few are particularly revolutionary.

Heat pumps could be much more efficient. They currently exchange at least three units of heat from inside and outside buildings per unit of electricity. When the thermodynamic effect of producing electricity is accounted for, the pumps become about as efficient as the best available gas furnaces. More than 11 million homes in southern and temperate climates in the United States already use them. Further savings could be made by improving the technologies used and by lowering the installation cost of pumps that extract heat from the ground. These can swap an impressive 8 units of heat for every unit of electricity. For water heating, we should encourage the adoption of electric and gas-fired pumps and solar heaters. Devices are now available that more than double the efficiency, at half the lifetime cost.

Meanwhile, liquid-desiccant technologies are starting to shift how people think about cooling and de-humidification. Current technology relies on the compression and expansion of vapour to drive heat out of a building. To remove moisture from the air, these often over-cool the air, then heat it up again. Solutions of brine, however, are much more efficient at extracting humidity. These systems use 30–50% less energy and are already commercially available from firms such as Advantix Systems in Sunrise, Florida, of which I am a vice-president. Systems that combine desiccant and evaporative cooling are on the horizon, and promise savings of 60–80%.

For lighting, simply increasing the use of efficient technologies such as compact fluorescent bulbs could lessen the energy used by 35%. The biggest opportunity lies with light-emitting diodes (LEDs), which could push energy use down by as much as 90%. The DOE awarded Amsterdam-based electronics firm Philips $10 million for producing an LED lamp that uses 85% less energy than an incandescent bulb, lasts roughly 25 times longer and provides the same quality of light. The bulb costs $50, but saves the consumer $200 in energy costs over its lifetime, compared with incandescent bulbs. The high purchase price is a sticking point, but manufacturers and the DOE plan to get the cost down to just $5 in the next 5–10 years.

"'Rebound' effects are are expected to soak up no more than 15% of energy savings."

Simple improvements to building insulation can also make a difference. For example, highly insulating windows currently cost about $1,000 more than standard windows for a typical home, with that difference earned back through energy savings in 5–10 years. But the profit differential isn't large enough for distributors to justify the time needed to convince customers to buy the product. One solution could be to create web tools or mobile apps that easily showcase the advantages of efficient products.

Savings of up to 20% or more are available through changes to maintenance and operations procedures. For example, 'learning thermostats' in homes assess the daily schedule and adjust the temperatures of

A heating and cooling system based on geothermal energy at a hospital in Greensburg, Kansas.

the rooms accordingly, and 'tune ups' on commercial buildings improve how they are run. Cars use control systems that are much more intelligent than those in our buildings.

INVEST TO SAVE

The barriers to 80% energy savings in buildings are well known. Consumers find it hard to evaluate the risks and benefits of technologies; investors may not benefit directly (why would a landlord invest in LED bulbs when tenants pay their own electricity bills?); and few of the latest technologies have been taken up by vendors. These small problems add up to big obstacles. But there are solutions.

Energy standards for US appliances were put in place between 1987 and 2010, for example, leading to savings of 3 quads in 2010. These are forecast to save a staggering 158 quads by 2070 (ref. 6). The DOE's Appliance and Equipment Standards programme is required by law to update appliance standards, but that takes effort. The most recent refrigerator standards, released in 2011, were developed after years of work, culminating in a 1,000-page report. Despite offering a return on investment of about a thousand to one, the rate of issuing standards has waxed and waned over time, and the programme has not always received the funding it needs. It is now on track, but no one knows for how long.

Building codes that were put in place in the 1970s have reduced energy use in newly constructed buildings by 30–40%, leading to a saving of 0.45 quads per year[7]. Codes developed in 2010 will soon be in place that should save an extra 1.25 quads per year by 2030 — a huge improvement — and others

in development could double those savings[7]. However, every state will need to adopt these codes individually, and many will need help with compliance.

Labels can be powerful, too. Since the DOE and the US Environmental Protection Agency introduced the ENERGY STAR programme in 1992 to test and label the efficiency of household equipment, energy use has fallen by about 1.6 quads per year[8]. Like appliance standards, this programme needs to be updated regularly. As new products raise the bar, old products need to lose their status. A recently introduced 'most efficient' ENERGY STAR label is a step in the right direction.

Further investment in R&D could make a difference. A 2001 analysis[9] of selected DOE-funded research projects showed net energy cost savings of around $30 billion from an investment of about $0.4 billion in energy-efficiency R&D: a return of 75 to 1.

Despite such rewards, the energy sector is hugely underfunded. Total R&D investments from US public and private sources come to just 0.4% of the amount spent on energy bills (about $1.5 billion per year in buildings-related R&D). This percentage is significantly lower than that spent in the pharmaceutical (20.5%) and aerospace and defence (11.5%) sectors[10]. Assuming the same return of 75 to 1, $40 billion would need to be invested to generate a cumulative $3 trillion of savings. Even if only $3 billion were invested per year until 2030, that would still be only about 1% of the money spent on buildings-related energy.
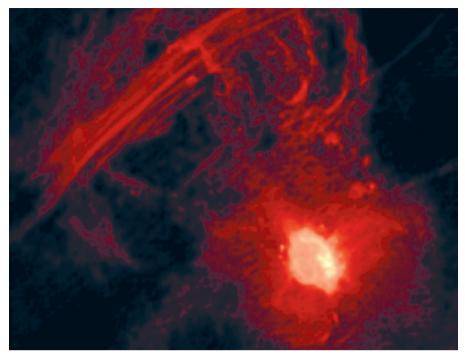
Strategic investment offers the opportunity to reduce the US annual energy bill by

hundreds of billions of dollars. Achieving an energy-efficient future is a cause that can benefit all of us. ∎

Philip Farese *is vice-president of sales support at Advantix Systems in Sunrise, Florida, the leading manufacturer of liquid-desiccant air conditioners. He was a senior energy analyst at the US National Renewable Energy Laboratory until early August 2012.*
e-mail: philip.farese@advantixsystems.com

1. US Energy Information Administration *Annual Energy Outlook 2012* (Department of Energy, 2012).
2. Farese, P., Gelman, R. & Hendron, R. *A Tool to Prioritize Energy Efficiency Investments* (National Renewable Energy Laboratory, in the press).
3. *Real Prospects for Energy Efficiency in the United States* (National Academies Press, 2009).
4. Jenkins, J., Nordhaus, T. & Shellenberger, M. *Energy Emergence: Rebound & Backfire as Emergent Phenomena* (Breakthrough Institute, 2011).
5. Sorrell, S., Dimitropoulos, J. & Sommerville, M. *Energy Policy* **37,** 1356–1371(2009).
6. Meyers, S., Williams, A. & Chan, P. *Energy and Economic Impacts of U.S. Federal Energy and Water Conservation Standards Adopted From 1987 Through 2010* (Lawrence Berkeley National Laboratory, 2012).
7. Belzer, D. & Cort, K. A. *Projected Impacts of the Building Energy Codes Program, 2012–2030* (Pacific Northwest National Laboratory, 2012).
8. Sanchez, M., Brown, R. E., Homan, G. K. & Webber, C. A. *2008 Status Report Savings Estimates for the ENERGY STAR Voluntary Labeling Program* (Lawrence Berkeley National Laboratory, 2007).
9. National Research Council *Energy Research at DOE: Was It Worth It? Energy Efficiency and Fossil Energy Research 1978 to 2000* (National Academies Press, 2001).
10. Holliday, C. *et al. Catalyzing American Ingenuity* (American Innovation Energy Council, 2010).

Radiowave images show filamentary structures near the black hole at the centre of our Galaxy.

COSMOLOGY

Enigma at the heart of the Universe

Mario Livio enjoys a tour of modern astronomy that speculates on how black holes drive cosmic evolution.

"The road of life twists and turns and no two directions are ever the same. Yet our lessons come from the journey, not the destination." So wrote US novelist Don Williams Jr. Caleb Scharf's engaging *Gravity's Engines* demonstrates the truth in this statement.

Scharf speculates that black holes rule everything in the cosmic landscape — from the large-scale structure of the Universe to life. Using rich language and a brilliant command of metaphor, he takes on some of the most intricate topics in theoretical and observational astronomical research. He weaves a wonderfully detailed tapestry of what modern astronomy is all about, from the complexities of cosmic microwave background studies to the X-ray mapping of galaxy clusters.

Scharf begins with the basics: an excellent overview of the history of astronomy and astrophysics, and illuminating examples to explain how the complete cosmic picture is deciphered. His discussion of how gravity shapes the Universe and determines its evolution is admirably clear, and covers the early history of X-ray and radio astronomy, as well as some of the perplexing objects (such as compact X-ray sources) discovered using new observational techniques.

He handles the complex physical processes that occur in the vicinity of black holes with aplomb and attention to detail. And he gets to the crux of the feedback process between the growth of mass-accreting black holes and the evolution of stellar bulges in their host galaxies.

I feel that Scharf begins to overstate his argument in his attempts to solidify the case

Gravity's Engines: How Bubble-Blowing Black Holes Rule Galaxies, Stars, and Life in the Cosmos
CALEB SCHARF
Scientific American:
2012. 272 pp.
$26/£18.99

for supermassive black holes as "regulators" of star formation in galaxies. There is no doubt that central black holes and the stellar bulges that surround them (when they exist) co-evolve. But the outer disks of spiral galaxies such as Andromeda are hardly affected at all by the presence of supermassive black holes.

I have quibbles, too, with the passages in which Scharf attempts to support his argument that black holes are the main engines driving everything from re-ionization and cosmic star formation to galactic evolution and the emergence of life. For instance, he argues that stellar-mass black holes played a crucial part in the cosmic phase transition associated with the re-ionization of the intergalactic medium. But the jury is still out on this particular process. The main agent of re-ionization may be ultraviolet photons from galaxies — a point that looks likely to be verified soon through observation.

Scharf also uses the story of our own Galaxy, the Milky Way, and statistical information collected through the citizen-science project Galaxy Zoo, to depict the "duty cycles" of activity and calm in the process of accretion onto supermassive black holes. He appeals to these to circumvent the apparent discrepancy between the idea of black holes as regulators and the fact that the Milky Way's central black hole is currently inactive.

All this culminates with his intriguing statement that "the entire pathway leading to you and me would be different or even nonexistent without the coevolution of galaxies with supermassive black holes and the extraordinary regulation they perform". Scharf admits that many steps remain uncertain and that numerous questions have yet to be answered. But he proposes that because the cosmic and galactic environments leading to the rise of complexity and life are part of black holes' galactic evolution, it is reasonable to ask what special things link us directly to that history.

However, I feel less certain than Scharf about the answer. He explains that the tight correlation between the masses of supermassive black holes and of stellar bulges at galaxies' centres reveals a co-evolution. It is equally certain that feedback from supermassive black holes had an important role in the ensuing star-formation history in the bulges of galaxies. But was this the key factor in determining whether life-bearing planets should exist or not? I doubt it. Still, the idea makes for a very interesting journey. ∎

Mario Livio *is an astrophysicist at the Space Telescope Science Institute in Baltimore, Maryland. His blog, A Curious Mind, explores science, art and the links between them. His upcoming book is* Brilliant Blunders, *to be published in May 2013. e-mail: mlivio@stsci.edu*

TECHNOLOGY

Pre–digital dreams

Josie Glausiusz strolls through an evanescent 'cabinet of wonders' exploring the surreal side of technology.

I flung myself onto the floor cushions inside Stan VanDerBeek's *Movie-Drome* and gazed at the flickering images projected on its curved ceiling. An attack of insomnia the night before gave me a good excuse to lie down and relax as pictures spun by: a dancing skeleton, a fork, a statue, a fried egg, a terrified face, the Statue of Liberty, the Kremlin, the words 'WHO WHAT HOW'. It seemed as if I was inhabiting a frenetic early version of the Internet — which, in a way, I was.

VanDerBeek, an experimental film-maker, constructed the original 9-metre-high movie-drome inside a grain silo in 1963. He foresaw a global network of such pods as "hubs for the distribution of knowledge via the universal language of the information age", connected by a network of satellites, television and telephones. He died in 1984, having never realized his dream. The reconstructed pod — a piece of this World Wide Web prototype — forms the centrepiece of *Ghosts in the Machine*, an exhibition at the New Museum of Contemporary Art in New York.

The exhibition is conceived as a "cabinet of wonders" encapsulating a prehistory of the digital age. It is designed to explore "the dreams and nightmares of technology" as expressed by more than 70 artists, writers and other visionaries. The eclectic jumble of installations conveys both our dread and our embrace of technology, with its power to transform life and death.

The most terrifying is a reconstruction — commissioned for a 1975 exhibition — of the *Harrow*. This fictional execution machine appears in Franz Kafka's short story *In the Penal Colony* (1919). In Kafka's tale, a person accused of an unspecified crime is strapped to a bed while an array of long needles carves the purported crime into their flesh.

With its cogs and gears, dangling wires and sharp needles suspended over a coffin-like bed, the wooden contraption evokes the terror that must face every prisoner confronted with today's electroshock devices and other torture technologies. Incongruously, the *Harrow* shares a space with Jeff Koons's vacuum cleaners, the *New Hoover Convertible Doubledecker*, encased in Plexiglass and lit with fluorescent bulbs. Koons's

piece is perhaps there to evoke the torment of domestic drudgery.

Another misconceived machine in the show, created by psychoanalyst Wilhelm Reich in 1940, is the Orgone Energy Accumulator, a wooden cubicle lined with sheet metal and steel wool. A follower of Freud, Reich believed that a concentration of "orgone", which he conceived of as a primordial form of energy, would increase "orgastic potency". This would, in turn, relieve

Robert Breer's *Floats* (1970) move imperceptibly about the gallery.

pent-up psychic and physical tensions, promoting general health and vitality, and even playing a part in curing cancer. Reich offered the box for testing to Albert Einstein, who pronounced it a dud.

Among those treated with orgone therapy was British computer-science pioneer Alan Turing, who was forced into treatment to "cure" his homosexuality. His proposed Turing Machine would simulate the behaviour of any other machine, including itself. Two years after being prosecuted for gross

indecency in 1952, Turing committed suicide, possibly by eating a cyanide-laced apple.

Henrik Oleson's *Some Illustrations to the Life of Alan Turing* (2008) appears in the same room as Reich's accumulator. This photo-collage series representing the dualities of Turing's life — public and private, mind and body, human and machine — depict Turing's face obscured by, among other objects, a large screw, a grid of zeroes and ones, and a half-eaten apple.

The machine as medicine shows up again in a piece of art constructed by Emery Blagdon, a sometime vagabond and tinkerer who once built a working tractor from the ground up. Born in 1907, Blagdon created intricately woven wire-and-bead chandelier-like mobiles, which he called "healing machines", and hung them in an outhouse on his Nebraska farm. He re-routed the power lines from his house into the shed, hoping to harness the electricity's healing powers. However, Blagdon died of cancer in 1986, the energy from his machines having failed to heal him.

There is much to admire in *Ghosts in the Machine*. Thomas Bayrle's 1989 *Madonna Mercedes*, for instance, is a mother-and-child collage that, on close inspection, is made up entirely of drawings of Mercedes-Benz automobiles. Günther Uecker's *New York Dancer IV* (1965), a nail-studded phallic sculpture 2 metres tall and powered by an electric motor, is more mystifying.

Mystification, however, can also be a thing of beauty and delight, as evinced by Gianni Colombo's *Elastic Space*, a three-dimensional grid made of illuminated, moving rubber bands. Designed to "catalyze variable perceptual experiences", the piece is mesmerizing and soothing.

I stood in the quiet, darkened room watching the glowing rubber bands slowly moving back and forth, up and down. I thought perhaps I would stay for a while. I could contemplate infinity. And maybe go to sleep. ∎

Josie Glausiusz *is a journalist based in New York City.*
e-mail: josiegz@gmail.com

Q&A Maja Matarić
The social roboticist

Maja Matarić, a computer scientist and neuroscientist at the University of Southern California in Los Angeles, makes robots that assist people with disabilities, children with autism and elderly people — a phenomenon explored in the film Robot and Frank (2012). *On the eve of its release, she talks about the future of socially assistive machines.*

How does *Robot and Frank* portray the human–robot interface?
The film asks whether people can relate to machines, and how robots can bridge the gap between what ageing people need and what their families can provide. In it, an elderly jewel thief convinces his caretaker robot to help him pull off a heist.

Did the film use a real robot?
The film-makers make a human in a robot suit look like a robot trying to look human. The robot's capabilities are necessarily beyond the current state-of-the-art: it serves as butler, nurse, chef and maid. Robots that manipulate objects in the real world are harder to engineer, and because these tasks require coordination and strength, safety is a major concern.

Are you designing those kinds of robots?
There's a whole realm of need that doesn't demand any physical work. Robots can help people to cope with depression, isolation, exercise and even social disorders such as autism. Most robotics is about getting the robot to do the right thing, whereas we make robots that get the person to do the right thing. Our socially assistive robots provide coaching, encouragement and companionship. We would never want to coerce people; rather, we aim to influence their behaviour. Someone who has had a stroke may need to spend hours each day working to regain function in an arm. A robot can use persuasion, motivational techniques and social manipulation to help the patient to persist with that. One student in my lab is studying how a robot might use the human quality of reciprocity. By helping a person with a task, such as shelving books, it would make it more likely that the person would adhere to the robot's requests, such as to do more exercise.

Can robots help with diagnosing autism?
With about one in every 100 children affected, early diagnosis is badly needed. We are designing a robot for diagnostic use in paediatric waiting rooms: as kids play with it, it could flag those with autism-like symptoms.

What about treatment of autism?
We have demonstrated that child-sized humanoid robots can encourage some children with autism to be more verbal and empathetic. The robots' life-like appearance and responsive behaviour seem to stimulate children to play with them and express empathy; when our robot didn't obey a command, one child said, "Now I know how my teacher feels." We're using a robot to model an autism behavioural therapy that aims to help a child gradually improve on a task. It will use verbal feedback and encouragement while playing games to encourage the child to imitate its poses and gestures. Ultimately, we hope to see the results of this research in affordable products for home use.

How can robots motivate people?
In a study we used Bandit, a child-sized humanoid robot on wheels, as an exercise coach. A person was seated across from Bandit and they played 'chair aerobics'. Sometimes Bandit led the exercises, and sometimes the user taught exercises to Bandit — many people love to be in charge some of the time. In another game, Bandit gave ever-longer exercise sequences to test memory and stamina. Through the use of computer vision, Bandit was able to detect the person's behaviour and performance, and respond appropriately by providing encouragement, increasing the levels of challenge and making humorous comments. We found that users strongly preferred Bandit to a computer simulation with the same games and behaviours, rating the robot as knowledgeable and caring. It effectively motivated people to exercise consistently.

Will caretaking robots take jobs from people?
Thirty years from now, there will be about as many old people as children and working-aged people in the United States. The population demographics are already creating a gap in care. Even if we expanded the immigration rate, we couldn't afford to fully train and pay everyone to be caregivers for the entire elderly population. Technology, including robotics, will have to be part of the solution.

Do assistive robots pose ethical issues?
I am concerned about claims that salespeople might make about these machines' capabilities. We don't want snake-oil promises of health benefits that aren't clinically validated. Other issues could arise from companies failing to advertise necessary software upgrades or implement recalls. And, as with people, when you put robots in the 'wild', their behaviour is not entirely predictable. If robots are continually learning and adapting to their users, it will be interesting to see the relationship dynamics that emerge between the two.

How can long-term robot–human relationships be managed?
We don't want attachments between humans and their robots to become dysfunctional. As with human–human relationships, it is best that people do not become overly dependent, submissive or abusive. Some ask what will happen if a well-loved robot breaks down. Its memory will be backed up and, for a price, an identical body can be built for its data to live in. The real question is, what is the warranty? ∎

INTERVIEW BY JASCHA HOFFMAN

Correspondence

Keep politics out of academia in Israel

Israel's thriving academic establishment is being ripped apart by the Israeli government's apparent drive to use academics for political ends.

To bolster its agenda of developing settlements in the West Bank occupied territories, the Israeli military's Council for Higher Education in Judea and Samaria has upgraded a small college there, formerly known as Ariel College, to university status. Millions of dollars are to be poured into the university's budget to attract researchers, circumventing Israel's usual channels for academic funding.

In response, more than 1,100 senior faculty members from all over Israel — over one-quarter of the country's total — have signed a petition rejecting the upgrade (see go.nature.com/dqte9c). The planned upgrade is also opposed by the presidents of all seven research universities in Israel and by the influential planning and budgeting committee of the state's Council for Higher Education.

The legitimacy of building settlements on occupied land is disputed in Israel and deplored worldwide. Ever since the occupation of the West Bank 45 years ago, Israeli academia has largely remained outside the political struggle. But Ariel University, which is closed to Palestinians, creates a dilemma for academics there: participate and thereby endorse the settlements, or boycott and be branded as the enemy. Such political meddling can only ruin Israel's academic structure in the long run.
Nir Gov *Weizmann Institute of Science, Rehovot, Israel.*
nir.gov@weizmann.ac.il

Turkey must build on science reforms

As chair of the Committee on Freedom and Responsibility in the Conduct of Science at the International Council for Science (ICSU), I was alarmed to read of the arrest in Turkey in June of Kemal Gürüz (http://doi.org/h47), twice a victim of the country's apparent crackdown on academic freedom (*Nature* **487**, 19–20; 2012). The committee monitors such cases and calls for greater protection of the human rights of individual scientists.

Gürüz, as a former president of Turkey's Council of Higher Education, has helped to raise education standards significantly in the country. He has also facilitated higher-education opportunities abroad for hundreds of Turkish students — particularly in the United States, where he was a Fulbright scholar and, later, chairman of the Turkish Fulbright Commission. His rearrest seems to relate to his attempted modernization of the education system, including provisions to educate Turkish students to become scientists.

Turkey's scientific contribution will depend on building up its research structure and continuing the reforms initiated by influential academics such as Gürüz.
Bengt Gustafsson *ICSU, Paris, France.*
bengt.gustafsson@fysast.uu.se

Time to accept conservation triage

Like the troubled California condors *Gymnogyps californianus* (*Nature* **486**, 451; 2012), more than 80% of endangered US species are imperilled by threats that cannot be eliminated, only managed. These species are "conservation reliant" (J. M. Scott *et al. Conserv. Lett.* **3**, 91–97; 2010).

For example, the endangered Kirtland's warbler (*Setophaga kirtlandii*) has exceeded its recovery goals as a result of control of the brown-headed cowbird (*Molothrus ater*), which lays its eggs in warbler nests, and maintenance of warbler habitat through vegetation management. These conservation actions will be required in perpetuity if the species is to avoid extinction. But funding of such management is inadequate for all but the most iconic species, such as the condor.

Society must commit to increased investment in conservation intervention, which will take greater political motivation and broader conservation partnerships than at present. We must begin to prioritize conservation efforts and accept that conservation triage may sometimes be necessary.
John A. Wiens *PRBO Conservation Science, Corvallis, Oregon, USA. jwiens@prbo.org*
Dale D. Goble, J. Michael Scott *University of Idaho, Moscow, Idaho, USA.*

Call for clear policy on deceased authors

William Laurance and his colleagues (*Nature* http://doi.org/h5f; 2012) identify four of their paper's authors as deceased and confirm that all had previously verified that they wanted to be co-authors. Without such indication, there is a risk that deceased authors might be included on questionable grounds — for example, as little more than a dedication.

Increasingly, authors' contributions to papers have to be explicitly justified. Although opinions of what constitutes enough to warrant authorship can vary between individuals and journals, most would endorse your journal's policy that "all the listed authors [must] have agreed all of the contents" of the final version of the submitted paper (see go.nature.com/qnanep).

Because this criterion may not be fulfilled for a deceased author who has genuinely contributed to the work, journals should not only denote authors who died before submission but also specify the date of their death and the contribution they made (see, for example, go.nature.com/pymlxb and go.nature.com/jxqr8n). The author list should not include unwarranted names of the deceased, for whom a dedication can be included in the acknowledgements.
Vincent Nijman *Oxford Brookes University, UK.*
vnijman@brookes.ac.uk

How to allay fears over nanomaterials

Regulating the use and engineering of nanomaterials is problematic owing to scientific uncertainty and controversy over their safety. To pave the way for regulatory approval and to build consumer confidence, the European Commission's Joint Research Centre and the European Academies Science Advisory Council have identified factors to help assess the benefits and risks of nanomaterials to health (see go.nature.com/utmi3d).

Our analysis indicates that formulation of a coherent public policy will depend on scientists closing knowledge gaps in safety research, on gathering more data to connect science and regulation, and on training graduate students in nanotechnology research. Policies will need to be flexible to accommodate fresh discoveries in this rapidly advancing technology.

More dialogue on nanomaterials is needed among private and public research groups, regulators and the public.

Genetic modification of crops has taught us that consumer acceptance as well as regulatory approval are needed for successful implementation of new bioscience technologies. To limit alarmist media assertions about nanomaterials, scientists have a responsibility to provide accessible and accurate information for the public.
Robin Fears *European Academies Science Advisory Council, Halle, Germany.*
Peter Gehr *University of Bern, Switzerland. gehr@ana.unibe.ch*
Elke Anklam *Institute for Health and Consumer Protection, EC Joint Research Centre, Ispra, Italy.*

TIGE IS THE MAN

An old-fashioned night out.

BY SUE LANGE

Tige is the man and he knows it. Mr Personality. Sharp-tie-wearing, punchline-busting, girl-sweet-talking. Everybody loves him; he puts up such a good front. And his parties? Sublime.

Tonight he's gone all out, giving the cats something they've never tasted. A trip into the past when things were dicey and anything could happen. Not like today where you know what's going to happen even before it does.

The guests arrived at Tige's in twos, couples holding tight as if in love. As if someday they'd reveal themselves to their one and only. Expose their naked feelings to their significant other late at night, maybe, when the curtains were closed and the children were in bed.

He watched them and wondered and laughed to himself. He wasn't even drunk yet, but he laughed, because he knew that before the night was through all would be revealed.

He'd spiked the rum punch with floating hothouse gardenias. He'd padded each ash-tray with green marijuana spliffs. His house delivery system was set to a groovalicious pop rock. The beat, key and bassline changed with each new arrival at the door. As the guests entered, they let their guard down for half a mo' so that the house could read their minds and respond with the appropriate musical vibe. The evening was electric.

At ten, the games began.

"Let us hark back to the age of chivalry," Tige said from his post in front of the defunct fireplace in the wall of his condo. He had a malicious grin on his face when he said: "When goils were goils, and men were men."

Everyone knew the bit of mass culture he was referring to. They knew it all by heart. And if they didn't, they could check it remotely, instantaneously.

Tige continued. "I'm talking about the late twentieth, ladies and gents. Back when people had open minds. Privacy was their birthright. It was easy to be a man or a goil back then."

They laughed again.

They didn't really know why. But the words sounded funny so they laughed.

"They had party games like we do, but they were different," Tige said. "They played Charades and Pin the Tail on the Donkey instead of our games like My 15 Minutes or Pass the Vid. There was one such oddity called Truth or Dare, and that's the one we'll hark back to tonight. It's all about daring to tell the truth. And daring to believe it. Remember boys and goils, it was a leap of faith back then. Back then, no one could check. Nowadays …" Tige moved his eyebrows up and down as if he was peeking under somebody's skirt.

"Und so!" he shouted with finger pointing to the ceiling. "Tonight we check!"

He waved his hands in the air, beckoning to the house and its mind-readingware installed in the walls. "The very latest in Sony Hive-Mind tech," he said, as if that was a statement, information they needed because they couldn't have guessed that Tige wouldn't be caught dead without the latest in anything.

"Tonight we make history," Tige declared. "For the first time in humanity's long flight from the trees, we dare to be truthful."

And with one final flourish, he passed his hand in front of his face and let down his guard.

The crowd of half-drunk, half-high, neo-hip-cons gasped. One young girl tittered as she saw something she'd never seen before: a big long thought. Tige remained steadfast even as his face turned a bit rose. He grew bolder and took in the faces of his guests, passing judgments as he scanned. They could see what he thought of them, who he liked and who he'd screwed and who he'd rather see dead.

They laughed and cried, guffawed and swore; he made some enemies in a quick five minutes. But his fears slowly left him and his conscience ran clear. A smile spread across his face. He felt fine.

"Come on in, my friends," he said. "It's great! A little draughty in the old noggin maybe, but in general I feel fine."

The crowd remained silent and thoughtful for a moment only and then they, not wanting to be left behind at the avant-garde event of the year, likewise passed their hands in front of their faces with a flourish. They let slip their shields and revealed what lay behind.

The house extracted the contents of their minds and published them in Tige's room. They pointed at each other and giggled with hands over their mouths as if trying to hold back thoughts that burbled out. Some turned scarlet; one man ran to the bathroom and cried. The more they drank, though, the more they howled. A good time was had by all.

At five in the morning, when dawn threatened to break, they sobered up and replaced the shields. They cleared their throats, adjusted their caps, straightened their skirts. One by one, they left alone. ∎

Sue Lange *is a writer living in Pennsylvania. Her parties are not to be missed.*

JACEY

CANCER

Exploiting collateral damage

Some mutations in tumour cells play no part in causing cancer, but they generate cellular weak spots that may allow tumour cells to be selectively killed by drugs. SEE ARTICLE P.337

BEN LEHNER & SOLIP PARK

The central challenge in anticancer therapy is to kill tumour cells without harming other cells of the body. A defining property of cancer cells is that they carry genetic changes that distinguish them from normal cells and that should, at least in theory, allow them to be specifically killed by drugs that target them. But in practice this represents a formidable problem. On page 337 of this issue, Muller et al.[1] outline a potential strategy that exploits 'passenger' mutations. These mutations can remove genetic redundancies in cancer cells, making them vulnerable and so facilitating their destruction.

Most cancer research focuses on the genetic mutations that cause cancer — 'driver' mutations that typically activate cancer-causing oncogenes or inactivate tumour-suppressor genes. Such mutations make logical drug targets, and, in recent decades, an explosion of knowledge about their mechanism of action has led to the development of several anticancer drugs. Examples include the successful inhibition of the proteins encoded by the oncogenes *BCR-ABL*, *BRAF* and *ERBB2* in patients with chronic myelogenous leukaemia, melanoma and breast cancer, respectively[2,3]. But in many other cancers it has proved difficult to target the driver mutations effectively, even though they have been known for a long time.

An alternative approach would be to target the vulnerabilities created by passenger mutations. Although the development of a tumour may depend on only a small number of driver mutations, tumour cells usually also contain other genetic changes that accumulate as the tumours develop, often through exposure to environmental mutagens such as tobacco smoke and ultraviolet light[4]. Solid tumours can contain tens of thousands of such mutations[4]. The key point is that these passenger mutations also distinguish cancer cells from normal cells. So if the presence of a passenger mutation means that a cell becomes sensitized to the effects of a particular drug, then this should allow the tumour cells to be specifically killed.

Muller et al. outline one way in which this might be achieved. Many essential processes in biology are performed by two or more genes
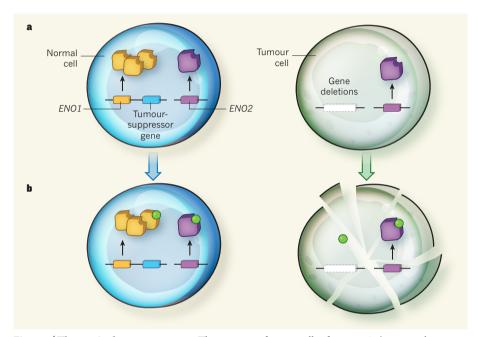
Figure 1 | Therapy in the passenger seat. The genomes of cancer cells often contain 'passenger' mutations, which are not directly related to the disease but which can create cellular vulnerabilities that allow the tumour cells to be killed by drugs targeting them. **a**, For example, Muller et al.[1] show that in the case of the brain cancer glioblastoma, a cancer-causing deletion of a tumour-suppressor gene also leads to loss of the *ENO1* gene, which encodes the enzyme enolase. This makes the tumour cells dependent on *ENO2*, which also encodes enolase but is expressed at lower levels than *ENO1* in brain cells. **b**, Normal cells will survive treatment with a drug that inhibits enolase activity (indicated by the green circle), because of their high expression of *ENO1*. The low levels of enolase in the tumour cells, however, mean that they will be killed by this treatment.

encoding proteins that have very similar biochemical functions. This 'partial redundancy' arises through the duplication of genes. It is a widespread phenomenon, and is maintained in genomes over vast evolutionary timescales[5]. One of the effects of partial redundancy is that the loss of one gene sometimes has little effect unless a second, related gene is also inhibited. Muller and colleagues' idea is that, if a passenger mutation removes or inactivates a duplicated gene, the tumour cells will become exquisitely sensitive to inhibition of the second, partially redundant gene (Fig. 1).

To test this idea, the authors hunted for passenger mutations that resulted in the deletion of a duplicated gene in a type of brain cancer called glioblastoma. They found that, in their tumour samples, the gene *ENO1* was deleted in several cases — most probably because it is located in a region of the genome that also

contains one or more tumour-suppressor genes that, when deleted, promote uncontrolled proliferation of the tumour cells. *ENO1* is one of three human genes that encode enolase, an enzyme required for the metabolic pathway that converts sugar to energy. These genes originated from duplications of an ancestral enolase gene more than 500 million years ago, and each gene is expressed to varying extents in different tissues. In the brain, for example, most enolase activity is provided by the ENO1 protein, with a smaller contribution from ENO2.

Consistent with their hypothesis, the authors found that glioblastoma cells in which *ENO1* is deleted are much more sensitive to inhibition of *ENO2* than other cells. They tested this sensitivity in cultured cells using two methods of inhibition — a known small-molecule inhibitor of enolase and a short hairpin RNA

molecule (shRNA) that binds specifically to enolase-encoding messenger RNAs and interferes with the production of enolase proteins. The authors found that in normal cells, because of the high levels of ENO1 protein present, ENO2-specific shRNA molecules or the enolase inhibitor had little effect. But in glioblastoma cells in which ENO1 was deleted, both treatments killed the cells (Fig. 1). Indeed, when the authors inhibited ENO2 activity in the ENO1-deficient glioblastoma cells, the cells could no longer form tumours when transferred into mice.

The study therefore provides a powerful proof of concept that the loss of redundancy in cancer genomes can be exploited to selectively kill tumour cells. In glioblastoma, the strategy works because the deleted duplicate gene (ENO1) normally provides most of the enolase activity, and there seems to be no compensatory increase in protein production from the second gene (ENO2). But compensatory changes in gene expression from duplicate genes may actually be common[6,7], and in other types of cancer the overall difference in protein activity between the tumour and normal cells may be much harder to target therapeutically. Moreover, the strategy must still be tested in animal models and in patients, particularly as a combination therapy with other drugs, because compensatory increases in the expression of the remaining gene duplicate may lead to the rapid development of drug resistance.

Finally, although not mentioned by Muller and colleagues, the vulnerabilities created by common passenger mutations are likely to extend far beyond the loss of protein redundancy[8]. Systematic screens to identify 'synthetic lethal' interactions[9,10], in which a drug or a gene inhibition kills only cells carrying a common passenger mutation, may be one way to identify these additional vulnerabilities. For instance, another recent study[11] presents an alternative strategy for targeting passenger mutations. It takes advantage of situations in which the activity of a single gene is reduced in cancer cells, for example through deletion of one of the two copies of the gene, rendering the cells highly sensitive to further inhibition of the same gene.

Muller and colleagues' results provide an impetus for two major conceptual advances: first, that the loss of protein redundancy provides a therapeutic opportunity to kill specific cells, and second, that passenger mutations may be the Achilles heel of cancer genomes. In our opinion, this second idea is the more important, and the plethora of passenger mutations present in cancer genomes should create many opportunities for personalized therapies. This underlines how crucial it is that, rather than simply focusing on cancer-causing genes and pathways, researchers also consider the therapeutic opportunities created by common passenger mutations. ∎

Ben Lehner and **Solip Park** are in the EMBL-CRG Systems Biology Unit, Centre for Genomic Regulation, Universitat Pompeu Fabra, 08003 Barcelona, Spain. **B. L.** is also at the Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona.
e-mail: ben.lehner@crg.eu

1. Muller, F. L. et al. Nature **488,** 337–342 (2012).
2. Sellers, W. R. Cell **147,** 26–31 (2011).
3. Workman, P. & de Bono, J. Curr. Opin. Pharmacol. **8,** 359–362 (2008).
4. Stratton, M. R., Campbell, P. J. & Futreal, P. A. Nature **458,** 719–724 (2009).
5. Vavouri, T., Semple, J. I. & Lehner, B. Trends Genet. **24,** 485–488 (2008).
6. Kafri, R., Bar-Even, A. & Pilpel, Y. Nature Genet. **37,** 295–299 (2005).
7. Burga, A., Casanueva, M. O. & Lehner, B. Nature **480,** 250–253 (2011).
8. Lehner, B. Trends Genet. **27,** 323–331 (2011).
9. Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Science **278,** 1064–1068 (1997).
10. Chan, D. A. & Giaccia, A. J. Nature Rev. Drug Discov. **10,** 351–364 (2011).
11. Nijhawan, D. et al. Cell http://dx.doi.org/10.1016/j.cell.2012.07.023 (2012).

APPLIED PHYSICS

Masers made easy

The technological potential of masers — the microwave equivalents of lasers — has been thwarted by their impractical operating conditions. A solid–state maser that works at room temperature may change that. SEE LETTER P.353

AHARON BLANK

Noise is an integral part of any electronic system, and it comes in many forms. We are all familiar with the noise that emerges from our radio when we try to listen to a distant station, or the 'snow-like' noise that appears on a detuned television screen. Scientists and engineers have been battling with noise for years, trying to make radio, television and mobile-phone communication better and clearer. In terms of noise reduction, however, the best performance is achieved only when electronic systems are operated at cryogenic temperatures. On page 353 of this issue, Oxborrow et al.[1] describe a solid-state microwave source and amplifier that has an exceptionally low noise level and that operates at room temperature. The device may find applications in space communication, radio astronomy and microwave spectroscopy.

Sources of noise can be roughly divided into those that are an intrinsic part of the signal of interest, about which not much can be done, and those that are added to the signal by external components, such as systems of electronic detection, reception or amplification. These extrinsic sources can be minimized, but the proper handling of signals and noise depends greatly on their characteristic frequency. One of the most crucial but challenging frequency bands for signal detection and amplification is the micro- and millimetre-wave range — frequencies of 1–100 gigahertz (1 gigahertz is 10^9 Hz). This part of the electromagnetic spectrum is commonly used for: surveillance radars; communication for mobile phones; space communication; radio astronomy (including the search for signals coming from extraterrestrial intelligence); and a variety of microwave-spectroscopy techniques.

In the first two of these applications, the noise levels that come with the signal are quite high. This means that low-cost amplifiers and detectors can be used at room temperature without significantly increasing the high intrinsic noise levels, and thus with almost no deterioration in the quality of the measured signals. However, in the other applications mentioned, the intrinsic noise can be very low. This means that considerable sensitivity to weak signals would be gained if the electronic detection and amplification of these signals were almost noiseless.

One of the earliest methods for amplifying faint signals in the microwave range, and still the best one, is based on the use of a maser — a device that performs 'microwave amplification by stimulated emission of radiation'. Masers are the predecessors and radio equivalents of lasers. They were at the centre of at least two Nobel prizes, one awarded to Charles Townes in 1964 for inventing the device, and the other to Arno Penzias and Robert Wilson in 1978 for the discovery of cosmic microwave background radiation — relic radiation from the Big Bang. The latter achievement was made possible only by the use of a maser amplifier whose noise contribution to the signal the researchers were detecting was so small that they had no choice but to 'blame' the background radiation for the noise in their system[2].

Masers of the type used in these experiments are based on crystals 'doped' with a paramagnetic substance — one that contains unpaired electron spins. In these systems, most of the unpaired electrons are found at the lowest energy level of the medium. But if the crystal is 'pumped' with microwave radiation, it is possible to reach a situation in which there are more of these electrons in an excited state than in a lower-energy state (a mechanism known

molecule (shRNA) that binds specifically to enolase-encoding messenger RNAs and interferes with the production of enolase proteins. The authors found that in normal cells, because of the high levels of ENO1 protein present, *ENO2*-specific shRNA molecules or the enolase inhibitor had little effect. But in glioblastoma cells in which *ENO1* was deleted, both treatments killed the cells (Fig. 1). Indeed, when the authors inhibited *ENO2* activity in the *ENO1*-deficient glioblastoma cells, the cells could no longer form tumours when transferred into mice.

The study therefore provides a powerful proof of concept that the loss of redundancy in cancer genomes can be exploited to selectively kill tumour cells. In glioblastoma, the strategy works because the deleted duplicate gene (*ENO1*) normally provides most of the enolase activity, and there seems to be no compensatory increase in protein production from the second gene (*ENO2*). But compensatory changes in gene expression from duplicate genes may actually be common[6,7], and in other types of cancer the overall difference in protein activity between the tumour and normal cells may be much harder to target therapeutically. Moreover, the strategy must still be tested in animal models and in patients, particularly as a combination therapy with other drugs, because compensatory increases in the expression of the remaining gene duplicate may lead to the rapid development of drug resistance.

Finally, although not mentioned by Muller and colleagues, the vulnerabilities created by common passenger mutations are likely to extend far beyond the loss of protein redundancy[8]. Systematic screens to identify 'synthetic lethal' interactions[9,10], in which a drug or a gene inhibition kills only cells carrying a common passenger mutation, may be one way to identify these additional vulnerabilities. For instance, another recent study[11] presents an alternative strategy for targeting passenger mutations. It takes advantage of situations in which the activity of a single gene is reduced in cancer cells, for example through deletion of one of the two copies of the gene, rendering the cells highly sensitive to further inhibition of the same gene.

Muller and colleagues' results provide an impetus for two major conceptual advances: first, that the loss of protein redundancy provides a therapeutic opportunity to kill specific cells, and second, that passenger mutations may be the Achilles heel of cancer genomes. In our opinion, this second idea is the more important, and the plethora of passenger mutations present in cancer genomes should create many opportunities for personalized therapies. This underlines how crucial it is that, rather than simply focusing on cancer-causing genes and pathways, researchers also consider the therapeutic opportunities created by common passenger mutations. ∎

Ben Lehner and **Solip Park** *are in the EMBL-CRG Systems Biology Unit, Centre for Genomic Regulation, Universitat Pompeu Fabra, 08003 Barcelona, Spain.* **B. L.** *is also at the Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona.*
e-mail: ben.lehner@crg.eu

1. Muller, F. L. *et al. Nature* **488,** 337–342 (2012).
2. Sellers, W. R. *Cell* **147,** 26–31 (2011).
3. Workman, P. & de Bono, J. *Curr. Opin. Pharmacol.* **8,** 359–362 (2008).
4. Stratton, M. R., Campbell, P. J. & Futreal, P. A. *Nature* **458,** 719–724 (2009).
5. Vavouri, T., Semple, J. I. & Lehner, B. *Trends Genet.* **24,** 485–488 (2008).
6. Kafri, R., Bar-Even, A. & Pilpel, Y. *Nature Genet.* **37,** 295–299 (2005).
7. Burga, A., Casanueva, M. O. & Lehner, B. *Nature* **480,** 250–253 (2011).
8. Lehner, B. *Trends Genet.* **27,** 323–331 (2011).
9. Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. *Science* **278,** 1064–1068 (1997).
10. Chan, D. A. & Giaccia, A. J. *Nature Rev. Drug Discov.* **10,** 351–364 (2011).
11. Nijhawan, D. *et al. Cell* http://dx.doi.org/10.1016/j.cell.2012.07.023 (2012).

APPLIED PHYSICS

Masers made easy

The technological potential of masers — the microwave equivalents of lasers — has been thwarted by their impractical operating conditions. A solid–state maser that works at room temperature may change that. SEE LETTER P.353

AHARON BLANK

Noise is an integral part of any electronic system, and it comes in many forms. We are all familiar with the noise that emerges from our radio when we try to listen to a distant station, or the 'snow-like' noise that appears on a detuned television screen. Scientists and engineers have been battling with noise for years, trying to make radio, television and mobile-phone communication better and clearer. In terms of noise reduction, however, the best performance is achieved only when electronic systems are operated at cryogenic temperatures. On page 353 of this issue, Oxborrow *et al.*[1] describe a solid-state microwave source and amplifier that has an exceptionally low noise level and that operates at room temperature. The device may find applications in space communication, radio astronomy and microwave spectroscopy.

Sources of noise can be roughly divided into those that are an intrinsic part of the signal of interest, about which not much can be done, and those that are added to the signal by external components, such as systems of electronic detection, reception or amplification. These extrinsic sources can be minimized, but the proper handling of signals and noise depends greatly on their characteristic frequency. One of the most crucial but challenging frequency bands for signal detection and amplification is the micro- and millimetre-wave range — frequencies of 1–100 gigahertz (1 gigahertz is 10^9 Hz). This part of the electromagnetic spectrum is commonly used for: surveillance radars; communication for mobile phones; space communication; radio astronomy (including the search for signals coming from extraterrestrial intelligence); and a variety of microwave-spectroscopy techniques.

In the first two of these applications, the noise levels that come with the signal are quite high. This means that low-cost amplifiers and detectors can be used at room temperature without significantly increasing the high intrinsic noise levels, and thus with almost no deterioration in the quality of the measured signals. However, in the other applications mentioned, the intrinsic noise can be very low. This means that considerable sensitivity to weak signals would be gained if the electronic detection and amplification of these signals were almost noiseless.

One of the earliest methods for amplifying faint signals in the microwave range, and still the best one, is based on the use of a maser — a device that performs 'microwave amplification by stimulated emission of radiation'. Masers are the predecessors and radio equivalents of lasers. They were at the centre of at least two Nobel prizes, one awarded to Charles Townes in 1964 for inventing the device, and the other to Arno Penzias and Robert Wilson in 1978 for the discovery of cosmic microwave background radiation — relic radiation from the Big Bang. The latter achievement was made possible only by the use of a maser amplifier whose noise contribution to the signal the researchers were detecting was so small that they had no choice but to 'blame' the background radiation for the noise in their system[2].

Masers of the type used in these experiments are based on crystals 'doped' with a paramagnetic substance — one that contains unpaired electron spins. In these systems, most of the unpaired electrons are found at the lowest energy level of the medium. But if the crystal is 'pumped' with microwave radiation, it is possible to reach a situation in which there are more of these electrons in an excited state than in a lower-energy state (a mechanism known

as population inversion) and to 'stimulate' them to decay into this lower state and emit radiation of the same frequency and phase as the incident radiation (Fig. 1a). In this way, the microwave radiation passing through the crystal is amplified. This emission and amplification is analogous to that of lasers, but occurs at much lower frequencies.

In such masers, the intrinsic noise, which is mainly caused by spontaneous photon emission, is very low — close to its physical limit. So similar designs are still used today in fundamental science, as well as in advanced space communications and radio-astronomy applications. However, for population inversion to be attained, the masers need to be cooled to liquid-helium temperatures (4.2 kelvin). This, together with their limited output power and the fact that their operation requires a rather large static magnetic field, has led to masers becoming a rare, almost extinct, species. Furthermore, in recent years, microwave technology has advanced so much that the noise levels achieved by cryogenically cooled, conventional semiconductor-based or superconducting amplifiers is comparable to that of masers, but with a much better power and bandwidth performance and far less physical complexity[3,4].

Now Oxborrow and colleagues have solved one of the most troubling aspects of solid-state maser operation. They demonstrate a device working at room temperature — a feat not previously achieved even though it has been five decades since the invention of the solid-state maser. They did this by using a special system: a crystal of the organic compound *p*-terphenyl doped with pentacene molecules. The energy levels of a pentacene molecule can be pumped optically to achieve large population inversion even at room temperature (Fig. 1b). A large population inversion, which leaves few electrons at the lower energy level, is required for the amplification process, and reduces the intrinsic noise of the maser.

In addition, Oxborrow *et al.* operated the maser using a microwave resonator that has low energy loss. Such low-loss resonators ensure that the device functions as an amplifier rather than as a sophisticated, but expensive, attenuator. In principle, their maser can be used to amplify weak microwave signals that have small intrinsic noise and without the need for any cooling. On the downside, like the 'old-fashioned' solid-state masers, its frequency of operation can be tuned only by an external magnetic field. However, this operational aspect is far less complicated than cryogenic cooling. Several other microwave devices, such as YIG oscillators, which also use external magnetic fields for frequency tuning, are in wide commercial use.

Finally, the authors' optically pumped maser delivers pulses of microwave radiation, rather than continuous signals. Ideally, amplifiers and sources of radiation should be able to operate
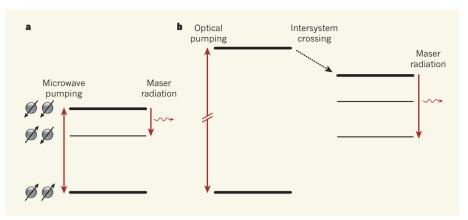

Figure 1 | Pumping masers. a, A conventional solid-state maser has at least three energy levels (horizontal lines), corresponding to the three possible spin combinations of two unpaired electrons (spheres; black arrows indicate spin states). Pumping the system's atomic medium with microwave radiation brings the electron spins from the lowest energy level to an excited state until the electron population of the levels is equalized. Subsequent decay of the electrons to the intermediate state results in the emission of maser radiation. The thickness of the lines denotes each level's population. **b**, In Oxborrow and colleagues' maser[1], electrons in pentacene molecules within a host crystal are optically pumped and undergo 'intersystem crossing', whereby they decay to a triplet state in which the uppermost level is the most populated. Decay from this level to the triplet's lowest level is accompanied by the emission of maser radiation.

in a continuous mode, which is more useful than 'part-time' operation. Nevertheless, with additional improvements — such as size reduction, a more efficient light-excitation protocol and the use of several resonators in parallel — the device may be useful for the applications mentioned previously. For example, the new maser can be seen as a first step towards a space-communication capability with which we could talk to extraterrestrials using 'phones' at room temperature. ∎

Aharon Blank *is in the Schulich Faculty of Chemistry, Technion-Israel Institute of Technology, Haifa 32000, Israel.*
e-mail: ab359@tx.technion.ac.il

1. Oxborrow, M., Breeze, J. D. & Alford, N. M. *Nature* **488**, 353–356 (2012).
2. Penzias, A. A. & Wilson, R. W. *Astrophys. J.* **142**, 419–421 (1965).
3. Schleeh, J. *IEEE Electron. Device Lett.* **33**, 664–666 (2012).
4. Eom, B. H., Day, P. K., LeDuc, H. G. & Zmuidzinas, J. *Nature Phys.* **8**, 623–627 (2012).

CELL BIOLOGY

Neither brown nor white

Fat cells are usually thought of as being either energy–storing white fat cells or food–burning brown fat cells. The identification of a third type of fat cell in mice and humans might open up new avenues for combating obesity.

BARBARA CANNON & JAN NEDERGAARD

In humans as in other mammals, fat is stored in white fat cells, whereas brown fat cells burn fat to generate heat. However, it seems that not everything is white or brown. Writing in *Cell*, Wu *et al.*[1] present convincing evidence that a distinct brown-like type of adipocyte (fat cell) exists within white fat depots.

Research on brown adipose tissue is developing rapidly. Until recently, it was thought that brown fat cells were merely modified white adipocytes. We now accept that the heat-producing, and thus food-consuming,

brown fat cells are in reality muscle relatives[2–4] (Fig. 1a). Both muscle cells and brown fat cells have many mitochondria — organelles often referred to as the cell's powerhouses — that burn food to liberate large amounts of energy. Muscle cells use this energy for the generation of ATP molecules that power muscle contraction. By contrast, brown adipocytes release the energy directly as heat, a uniquely mammalian ability gained through the development of the mitochondrial protein UCP1 (uncoupling protein 1). A deficiency of UCP1 leads not only to a lack of heat production but also to an increased propensity for

as population inversion) and to 'stimulate' them to decay into this lower state and emit radiation of the same frequency and phase as the incident radiation (Fig. 1a). In this way, the microwave radiation passing through the crystal is amplified. This emission and amplification is analogous to that of lasers, but occurs at much lower frequencies.

In such masers, the intrinsic noise, which is mainly caused by spontaneous photon emission, is very low — close to its physical limit. So similar designs are still used today in fundamental science, as well as in advanced space communications and radio-astronomy applications. However, for population inversion to be attained, the masers need to be cooled to liquid-helium temperatures (4.2 kelvin). This, together with their limited output power and the fact that their operation requires a rather large static magnetic field, has led to masers becoming a rare, almost extinct, species. Furthermore, in recent years, microwave technology has advanced so much that the noise levels achieved by cryogenically cooled, conventional semiconductor-based or superconducting amplifiers is comparable to that of masers, but with a much better power and bandwidth performance and far less physical complexity[3,4].

Now Oxborrow and colleagues have solved one of the most troubling aspects of solid-state maser operation. They demonstrate a device working at room temperature — a feat not previously achieved even though it has been five decades since the invention of the solid-state maser. They did this by using a special system: a crystal of the organic compound *p*-terphenyl doped with pentacene molecules. The energy levels of a pentacene molecule can be pumped optically to achieve large population inversion even at room temperature (Fig. 1b). A large population inversion, which leaves few electrons at the lower energy level, is required for the amplification process, and reduces the intrinsic noise of the maser.

In addition, Oxborrow *et al.* operated the maser using a microwave resonator that has low energy loss. Such low-loss resonators ensure that the device functions as an amplifier rather than as a sophisticated, but expensive, attenuator. In principle, their maser can be used to amplify weak microwave signals that have small intrinsic noise and without the need for any cooling. On the downside, like the 'old-fashioned' solid-state masers, its frequency of operation can be tuned only by an external magnetic field. However, this operational aspect is far less complicated than cryogenic cooling. Several other microwave devices, such as YIG oscillators, which also use external magnetic fields for frequency tuning, are in wide commercial use.

Finally, the authors' optically pumped maser delivers pulses of microwave radiation, rather than continuous signals. Ideally, amplifiers and sources of radiation should be able to operate

Figure 1 | Pumping masers. a, A conventional solid-state maser has at least three energy levels (horizontal lines), corresponding to the three possible spin combinations of two unpaired electrons (spheres; black arrows indicate spin states). Pumping the system's atomic medium with microwave radiation brings the electron spins from the lowest energy level to an excited state until the electron population of the levels is equalized. Subsequent decay of the electrons to the intermediate state results in the emission of maser radiation. The thickness of the lines denotes each level's population. **b**, In Oxborrow and colleagues' maser[1], electrons in pentacene molecules within a host crystal are optically pumped and undergo 'intersystem crossing', whereby they decay to a triplet state in which the uppermost level is the most populated. Decay from this level to the triplet's lowest level is accompanied by the emission of maser radiation.

in a continuous mode, which is more useful than 'part-time' operation. Nevertheless, with additional improvements — such as size reduction, a more efficient light-excitation protocol and the use of several resonators in parallel — the device may be useful for the applications mentioned previously. For example, the new maser can be seen as a first step towards a space-communication capability with which we could talk to extraterrestrials using 'phones' at room temperature. ∎

Aharon Blank *is in the Schulich Faculty of Chemistry, Technion-Israel Institute of Technology, Haifa 32000, Israel.*
e-mail: ab359@tx.technion.ac.il

1. Oxborrow, M., Breeze, J. D. & Alford, N. M. *Nature* **488,** 353–356 (2012).
2. Penzias, A. A. & Wilson, R. W. *Astrophys. J.* **142,** 419–421 (1965).
3. Schleeh, J. *IEEE Electron. Device Lett.* **33,** 664–666 (2012).
4. Eom, B. H., Day, P. K., LeDuc, H. G. & Zmuidzinas, J. *Nature Phys.* **8,** 623–627 (2012).

CELL BIOLOGY

Neither brown nor white

Fat cells are usually thought of as being either energy–storing white fat cells or food–burning brown fat cells. The identification of a third type of fat cell in mice and humans might open up new avenues for combating obesity.

BARBARA CANNON & JAN NEDERGAARD

In humans as in other mammals, fat is stored in white fat cells, whereas brown fat cells burn fat to generate heat. However, it seems that not everything is white or brown. Writing in *Cell*, Wu *et al.*[1] present convincing evidence that a distinct brown-like type of adipocyte (fat cell) exists within white fat depots.

Research on brown adipose tissue is developing rapidly. Until recently, it was thought that brown fat cells were merely modified white adipocytes. We now accept that the heat-producing, and thus food-consuming,

brown fat cells are in reality muscle relatives[2–4] (Fig. 1a). Both muscle cells and brown fat cells have many mitochondria — organelles often referred to as the cell's powerhouses — that burn food to liberate large amounts of energy. Muscle cells use this energy for the generation of ATP molecules that power muscle contraction. By contrast, brown adipocytes release the energy directly as heat, a uniquely mammalian ability gained through the development of the mitochondrial protein UCP1 (uncoupling protein 1). A deficiency of UCP1 leads not only to a lack of heat production but also to an increased propensity for

obesity, at least in mice[5]. And, importantly, although it was thought that in humans active brown fat existed only in babies and was progressively lost later in life, it is now clear that adults have active brown adipose tissue[6–11], and that the absence of this tissue might even be functionally linked to obesity[8].

When animals are chronically exposed to cold, the ability of brown fat to produce heat is enhanced manyfold, owing to an increase in the total amount of UCP1 in the tissue. In addition, certain apparently white adipose depots can undergo a shift in colour from whitish to brownish[12–14]. This 'browning' process is accompanied by a large increase in the expression of the UCP1 gene, relative to the very low starting level. A similar browning can be promoted in animals by various types of genetic manipulation. Earlier, when it was thought that brown fat cells were just some kind of modified white fat cell, the browning phenomenon seemed understandable. However, because we now know that brown fat cells are related to muscle and not to white fat cells, it is unlikely that the UCP1-expressing cells that appear in white fat after browning are classical brown fat cells. Indeed, previous studies[13–15] have suggested the existence of a subpopulation of cells in white fat depots that have the ability to express UCP1, and therefore to become brownish.

To explore whether such a subpopulation indeed exists, Wu et al. isolated progenitor cells (cells that can give rise to specialized cell types) from the subcutaneous white adipose tissue of mice and 'immortalized' these cells. Through this procedure, the cells acquired the ability to proliferate indefinitely and so yielded stable cell lines. Among hundreds of such cell lines, about 30 could be induced to accumulate fat.

The authors then analysed the gene-expression profiles of these cell lines. About a third of the cell lines showed an expression profile that was similar, although not identical, to that of brown adipocytes, whereas the rest displayed a quite different but rather homogeneous gene-expression pattern. Therefore, the original, apparently white, adipose tissue depot contained progenitors of genuine white fat cells, as well as a subpopulation of progenitors of a different type of adipocyte that — given an adequate stimulus — could express UCP1 and thus generate heat: the 'beige' or 'brite' (brown-in-white) cells (Fig. 1b). The nomenclature for these fascinating cells is presently unsettled, and they are also known as 'ectopic brown fat cells' or 'inducible brown fat cells'.

These results raise several questions. What is the origin of the two types of cell? Are there two distinct types of pericytes (cells that are wrapped along the blood capillaries and supposedly give rise to fat cells)? Or does one cell type derive, for example, from the bone marrow? How can certain white-like adipocytes, which in general possess very few mitochondria, suddenly enhance their mitochondrial complement during the 'beiging/britening' process?

Moreover, there is an additional twist to Wu and colleagues' story — and it concerns humans. The authors identified a group of mouse genes that were characteristically expressed in either beige/brite or brown fat cells. When they examined the gene-expression profile of brown fat from 11 adult humans, they unexpectedly found that the profile was closer to that of mouse beige/brite cells than to that of mouse brown cells. So, the brown fat of adult humans might not really be 'brown', but might instead be equivalent to mouse beige/brite adipocytes.

This suggestion is surprising in several ways. Brown fat in adult humans[6–11] is found in practically the same regions of the body as it occurs in babies. As described by Aherne and Hull in a classic paper[16] in 1966, brown adipose tissue in babies is present partly as a thin interscapular sheet between the shoulder blades (this sheet, as we now know, is not detectable in adults) but mainly as many smaller masses around the muscles and blood vessels of the neck — extensions of these masses pass under the clavicles to form rather large deposits in the axillae. Do Wu and colleagues' observations nevertheless mean that humans, even when we are babies, do not have genuine brown fat (that is, of the same type as that of other mammals)? Or is brown fat exchanged for beige/brite fat at some time during early human life? And is there a principal difference between the white (subcutaneous) adipose tissue in mice and humans? How should we modify our current mouse models to make them more similar to humans with respect to adipose tissues? And finally, can the beiging/britening process be stimulated so that we can obtain more food-burning cells that could help to combat obesity? ∎

Barbara Cannon and **Jan Nedergaard** *are at The Wenner-Gren Institute, Stockholm University, Stockholm SE-106 91, Sweden. e-mail: barbara.cannon@wgi.su.se*

1. Wu, J. *et al. Cell* **150,** 366–376 (2012).
2. Timmons, J. A. *et al. Proc. Natl Acad. Sci. USA* **104,** 4401–4406 (2007).
3. Seale, P. *et al. Nature* **454,** 961–967 (2008).
4. Cannon, B. & Nedergaard, J. *Nature* **454,** 947–948 (2008).
5. Feldmann, H. M., Golozoubova, V., Cannon, B. & Nedergaard, J. *Cell Metab.* **9,** 203–209 (2009).
6. Nedergaard, J., Bengtsson, T. & Cannon, B. *Am. J. Physiol.* **293,** E444–E452 (2007).
7. Cypess, A. M. *et al. N. Engl. J. Med.* **360,** 1509–1517 (2009).
8. Saito, M. *et al. Diabetes* **58,** 1526–1531 (2009).
9. van Marken Lichtenbelt, W. D. *et al. N. Engl. J. Med.* **360,** 1500–1508 (2009).
10. Virtanen, K. A. *et al. N. Engl. J. Med.* **360,** 1518–1525 (2009).
11. Zingaretti, M. C. *et al. FASEB J.* **23,** 3113–3120 (2009).
12. Young, P., Arch, J. & Ashwell, M. *FEBS Lett.* **167,** 10–14 (1984).
13. Cinti, S. *et al. J. Histochem. Cytochem.* **50,** 21–31 (2002).
14. Guerra, C. *et al. J. Clin. Invest.* **102,** 412–420 (1998).
15. Petrovic, N. *et al. J. Biol. Chem.* **285,** 7153–7164 (2010).
16. Aherne, W. & Hull, D. *J. Pathol. Bacteriol.* **91,** 223–234 (1966).
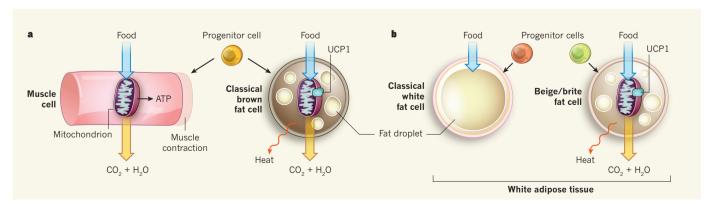
Figure 1 | Different progenitors, different colours. a, Cells that make up muscle and brown fat are derived from the same progenitor cell. Muscle cells and classical brown fat cells contain many mitochondria, organelles in which food is burned to produce carbon dioxide and water. The energy released is harvested in the form of ATP molecules that support contraction in muscles, whereas it is directly released as heat in brown fat. This heat generation requires the presence of the UCP1 protein. **b,** Wu et al.[1] identified two types of progenitor cell in mouse white adipose tissue. These cells give rise to classical white fat cells (which store energy in the form of fat droplets) and to 'beige' or 'brite' fat cells. Beige/brite cells express UCP1 and are functionally similar to brown cells, in spite of being derived from different progenitors.

ASTROPHYSICS

One of the first of the second stars

The chemical content of a star that was born relatively shortly after the formation of the Milky Way calls into question conventional understanding of how stars formed in the early Universe.

JOHN COWAN

How did the first stars form early in the history of our Galaxy and the Universe, and what were these stars like? The Big Bang produced only hydrogen, helium and some lithium, so the first stars would have contained only those elements. Given that the stars became extinct long ago, they were probably quite massive — the more massive a star (with masses more than 80 times that of the Sun a possibility), the shorter its lifetime (as little as a few million years). However, we can learn much about the early conditions in our Galaxy, including the types of element formed and the nature of these first stars, by studying surviving stars from a second generation of stars, which formed from the debris of the first. These surviving 'second stars', which are less massive than the Sun and live for many billions of years, can be found in the Galactic halo. Writing in *Astronomy & Astrophysics*, Caffau et al.[1] describe a Galactic halo star that is one of the earliest members of this second generation.

This new work is an extension, and expansion, of Caffau and colleagues' initial report[2] on this star, which is catalogued in the Sloan Digital Sky Survey and known by the non-sexy name of SDSS J102915+172927. The authors used the Very Large Telescope (VLT) in Chile to undertake a detailed spectroscopic analysis[1] of the elements contained in the star, which is in the centre of the constellation Leo. They report that the star has an extremely low abundance of iron — approximately 1/130,000 that of the Sun. Astronomers refer to a star's iron abundance as its metallicity. This metallicity generally correlates with Galactic time: iron is produced in exploding stars, known as supernovae, and its abundance in our Galaxy has increased progressively over time. Thus, the most iron-deficient — or extremely metal-poor (EMP) — stars are among the oldest, having formed relatively soon after the Galaxy itself, early in the history of the Universe.

And yet SDSS J102915+172927 does not have the lowest known iron abundance. Two other stars[3,4] are even more iron-poor, implying that they are older than the star in Leo. Compared to the Sun, however, both stars have a large abundance of carbon and nitrogen relative to iron, as has been noted in several other EMP stars. By contrast, Caffau et al.[1] found that carbon and nitrogen were not enhanced in SDSS J102915+172927 — the abundance of these elements with respect to that of iron is consistent with solar values. Similarly to some other EMP stars[4], the abundance of oxygen could not be measured in SDSS J102915+172927. However, Caffau and colleagues also defined a metal-mass fraction based on the total abundance of all of the elements heavier than helium, rather than just that of iron. They report that SDSS J102915+172927 has the lowest such value ever measured, and argue that this makes its composition similar to that of the primordial gas that existed shortly after the Big Bang.

The types and abundances of these elements in old, low-metallicity stars are crucial to our understanding of what happened before the stars' formation. First, it is a puzzle how a low-mass star such as SDSS J102915+172927 (which is less massive than the Sun) even formed early in the history of the Galaxy at a time when high-mass stars would seem to be more common. Observational and theoretical studies[5,6] have suggested that elements such as carbon or oxygen are necessary for cooling (low-mass) parent gas clouds sufficiently for them to eventually collapse and form low-mass stars. The difference between SDSS J102915+172927, which has a relatively low carbon abundance, and carbon-enhanced metal-poor (CEMP) stars calls into question what is normal for these early stars[7]. But with so few stars of this type observed, it is difficult to discern a pattern.

Perhaps SDSS J102915+172927 is older than the other EMP stars observed: the relationship between iron abundance and time might not be entirely linear so early on, and thus the lowest metallicity star might not be the oldest. SDSS J102915+172927 might fall into the transition region between the first generation of stars (sometimes referred to as Population III) and the second generation, or Population II; halo, EMP and CEMP stars belong to the latter group.

Or perhaps SDSS J102915+172927 formed in a region of the Galaxy that had particularly low levels of elements heavier than helium.

Although carbon can be produced internally in ageing (giant) stars, this star is probably an (unevolved) main-sequence star — a stellar phase similar to the current state of the Sun — and not a giant or a sub-giant[2]. This means that the carbon and other heavy elements observed in SDSS J102915+172927 must have been synthesized in a supernova and then incorporated into the gas that would form new stars.

The chemical-abundance pattern observed in SDSS J102915+172927 is consistent with what is predicted for such a supernova event[2]. Yet this star in Leo does not have enhanced carbon, in sharp contrast to one of the unevolved carbon-enhanced EMP stars, which also has detectable strontium[4] — a rare heavy element made only in supernovae by a series of neutron captures.

A large scatter in the abundance of heavy, neutron-capture elements with respect to iron is observed for metal-poor stars. (There are many metal-poor halo stars but only a few EMP stars.) This scatter suggests that, at early times, the Galaxy was an unmixed, or non-homogeneous entity, with individual element-synthesis events — that is, supernovae — scattered throughout the halo[8]. Thus, the greatly varying carbon abundances in these early stars might also result from this heterogeneity. Clearly, observations of additional stars will be needed to probe this early phase of the Milky Way.

Further support for this early history of the Galaxy was provided by Caffau and colleagues' measurement[1,2] of the lithium abundance in SDSS J102915+172927. Lithium, produced in the Big Bang, is expected to be at a uniform and primordial abundance level (denoted as the Spite plateau[9]) in these early stars. Surprisingly, however, the measured lithium abundance in this, and another[4], EMP star is low, below the observed Spite plateau. This suggests, at least for some of these early stars, that there is probably a lithium-destruction mechanism occurring during star formation. It will be crucial to find additional evidence of variations in lithium abundance to better understand the earliest stages of star formation.

There will be more stars such as SDSS J102915+172927. Only a small fraction of the thousands of stars in the SDSS database has been observed with the VLT. Nevertheless, this new discovery is a valuable first step in filling in the gaps in our knowledge of the early history of the Universe, and of how stars and elements were formed. ∎

John Cowan *is in the Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma 73019, USA.*
e-mail: cowan@nhn.ou.edu

1. Caffau, E. *et al. Astron. Astrophys.* **542,** A51 (2012).
2. Caffau, E. *et al. Nature* **477,** 67–69 (2011).

3. Christlieb, N. et al. Nature 419, 904–906 (2002).
4. Frebel, A. et al. Nature 434, 871–873 (2005).
5. Bromm, V. & Larson, R. B. Annu. Rev. Astron. Astrophys. 42, 79–118 (2004).
6. Frebel, A., Johnson, J. L. & Bromm, V. Mon. Not. R. Astron. Soc. 380, L40–L44 (2007).
7. Norris, J. E. Nature 477, 37–39 (2011).
8. Sneden, C., Cowan, J. J. & Gallino, R. Annu. Rev. Astron. Astrophys. 46, 241–288 (2008).
9. Spite, M. & Spite, F. Nature 297, 483–485 (1982).

NEUROSCIENCE

Circuits drive cell diversity

Neurons of the same type can show functional differences. It turns out that this diversity is in part the result of the cells' adaptation to their specific neural networks. SEE LETTER P.375

NATHANIEL URBAN & SHREEJOY TRIPATHY

Modern manufacturing was revolutionized by the use of interchangeable parts so similar in their function that any one could effectively replace any other. Making such parts meant that manufacturers did not need to keep track of which nut worked with which bolt or which piston was intended to go in which cylinder. Similarly, most physiological analyses ignore cell-to-cell variation and focus instead on differences between cell types, as though each cell of a specific class were functionally equivalent to any other cell of the same type. However, neuroanatomists have long marvelled at the snowflake-like diversity apparent in the shapes of individual neurons, even within a cell type. And recent analyses have demonstrated that same-class neurons show substantial heterogeneity in their intrinsic properties[1–3], although the origin of such diversity is poorly understood. On page 375 of this issue, Angelo et al.[4] report that physiological variability among mitral cells (a type of neuron in the olfactory system) is at least partly caused by differences in the inputs that they receive.

Sensory neurons in the nose are activated when odorant molecules bind to specific receptor proteins on the neurons' surface. This activation is transmitted as an excitatory signal along nerve fibres (axons) that terminate in structures called glomeruli, which cover the surface of the brain's olfactory bulb. Each of about 2,000 glomeruli in the mouse olfactory bulb receives axons from sensory neurons that express a single type of odorant receptor (Fig. 1).

Mitral cells receive these excitatory signals from the sensory neurons and are the main source of the olfactory bulb's output to the cerebral cortex; of note, each mitral cell is connected to a single glomerulus. Although they have long been considered to be a single class of neuron, mitral cells show substantial variation in their physiological properties[1,2]. This is probably because of differences in the expression of certain ion channels — pores in the cell membrane that mediate the transmission of electrical signals by allowing the passage of specific ions. Indeed, Angelo et al. have previously observed[1] that mitral cells display considerable diversity in the magnitude of a transmembrane current known as a 'sag'. This

current, which is generated when the resting electric potential across the neuron's membrane is intensified, is believed to regulate the sensitivity and timing of mitral cells' responses to periodic sensory inputs (such as those generated when an animal sniffs rapidly).

To explore the causes of variability in sag currents, the authors simultaneously recorded the sag from pairs of mitral cells belonging to either the same or different glomeruli, in slices of mouse olfactory bulb. Although sag magnitude was highly variable from one cell to the next, the variability was markedly lower between mitral cells that received input from the same glomerulus than between cells that received input from different glomeruli.

Why are same-glomerulus mitral cells so homogeneous? The hypothesis explored by Angelo et al. is that neurons with similar levels of activity acquire similar physiological features. Pairs of mitral cells connected to the same glomerulus have the same sources of input (Fig. 1) and will therefore have more similar activity levels than random pairs of mitral cells. If the expression of sag-mediating ion channels is affected by a neuron's activity, then this could account for the relative homogeneity of mitral cells connected to the same glomerulus.

To address this issue, the authors cleverly
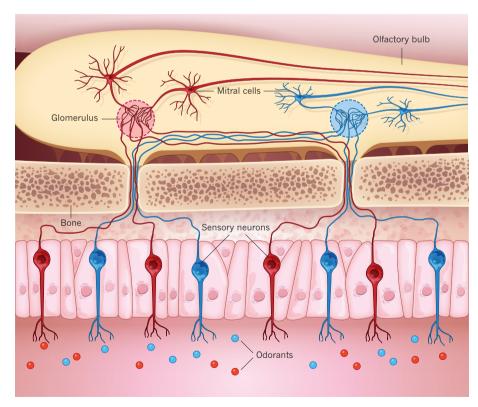
Figure 1 | The importance of network affiliation. Sensory neurons in the nose are activated by specific odorants or groups of odorants (different types are indicated by different colours), and transmit this information to mitral cells, another type of neuron in the brain's olfactory bulb. The transmission takes place in spherical structures called glomeruli, each of which receives input from a single type of sensory cell. Angelo et al.[4] report that mitral cells that are linked to the same glomeruli are functionally more similar to each other than to mitral cells that connect to other glomerular networks. This result suggests that input heterogeneity drives functional diversity of mitral cells.

50 YEARS AGO

A REPORT entitled *The Supply of Science Teachers in England and Wales* ... states that, immediately, there is need for an increase of between 28,000 and 34,000 over the present number of graduate science teachers in schools, technical colleges, training colleges and universities, which stands to-day at about 20,600 ... The shortage of science teachers is seen to be one part of the general teacher shortage, but many reasons are given for regarding it as a separate and specially urgent problem.
From *Nature* 18 August 1962

100 YEARS AGO

In 1904 Dr. J. R. Ashworth and I published ... observations on aged individuals of *Sagartia troglodytes* then and still in the possession of Miss Jessie Nelson in Edinburgh. After eight years these anemones are still in excellent health, having been in captivity for considerably more than half a century. In one respect I fear we did them an injustice, namely in attributing cannibalism to them, the error being probably due to the observation of the birth of young from a parent the tentacles of which were not fully expanded. Recently I chanced to notice a young Sagartia attached to a small piece of seaweed floating free in the aquarium. A slight agitation of the water was sufficient to bring the young anemone in contact with the tentacles of one of the patriarchs of its own species. They immediately closed round it and a small part of the disk became emarginate. The greater part, however, was not sensibly affected, and the mouth remained closed. In less than two minutes the folded-in tentacles uncurled and the young anemone was thrust away with some force ... Neither the young one nor the tentacles that held it were apparently affected in any way.
From *Nature* 15 August 1912

made use of an existing transgenic mouse strain, the 'monoclonal nose' mouse, in which almost all olfactory sensory neurons express the same odorant receptor. In these animals, most mitral cells receive similar inputs (regardless of their glomerular association) and would therefore be expected to show similar sag levels. Indeed, Angelo *et al.* found that randomly selected pairs of mitral cells in these mice showed much less diversity in sag magnitude than did equivalent pairs in control animals, consistent with the hypothesis that sag variation reflects heterogeneity in cellular activity.

One crucial question that the authors did not address is: what is the functional role, if any, of neuronal diversity? Theoretical studies have demonstrated that diversity could enhance information coding[2,5] and neural synchronization[6]; such studies provide valuable guidance for understanding how cellular diversity might support, or impair, brain function. This may tempt us to think that the robustness and power of brains as computational devices derive in part from the bespoke nature of individual neurons. However, a definitive answer to the question of whether neural variability is functionally important, or simply an unavoidable consequence of the imprecision of biological systems, awaits future experiments in which diversity can be controlled while probing behaviour. ■

Nathaniel Urban and **Shreejoy Tripathy** are in the Department of Biological Sciences and Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA.
e-mail: nurban@cmu.edu

1. Angelo, K. & Margrie, T. W. *Sci. Rep.* **1,** 50 (2011).
2. Padmanabhan, K. & Urban, N. N. *Nature Neurosci.* **13,** 1276–1282 (2010).
3. Marsat, G. & Maler, L. *J. Neurophysiol.* **104,** 2543–2555 (2010).
4. Angelo, K. *et al. Nature* **488,** 375–378 (2012).
5. Stocks, N. G. *Phys. Rev. Lett.* **84,** 2310–2313 (2000).
6. Longtin, A. & Mejias, J. F. *Phys. Rev. Lett.* **108,** 228102 (2012).

OCEANOGRAPHY

The trouble with the bubble

Past studies have suggested that the ocean's nitrogen budget has a deficit of fixed nitrogen. This view may now change, thanks to upward revisions of the rate of nitrogen input through biological activity. SEE LETTER P.361

ANGELICQUE E. WHITE

As is the case for gardens, forests and fields, the availability of fixed nitrogen (such as nitrate and ammonia) can limit the productivity of our seas. Oceanographers are therefore interested in the relative magnitudes of sources and sinks of fixed nitrogen as dynamic controls of ocean fertility. The trajectory of the oceanic nitrogen inventory has been a long-debated and contentious topic: direct rate measurements of biologically mediated nitrogen fluxes suggest that the ocean is currently being depleted of this resource, whereas geochemical evidence indicates a steady-state, balanced nitrogen budget[1]. On page 361 of this issue, Großkopf *et al.*[2] report that applying a new twist to an old method may reconcile these views and lead to a revision of our understanding of the present-day oceanic fixed-nitrogen budget*.

Unlike other elemental cycles, which are predominantly influenced by riverine, atmospheric or sedimentary fluxes (such as the iron and phosphorus cycles), the oceanic inventory of fixed nitrogen is largely set by the push and pull of two biological processes. First, fixed nitrogen is added to the oceans by microorganisms known as diazotrophs, which convert the nearly limitless supply of atmospheric nitrogen gas (N_2) dissolved in sea water to ammonia. Second, in sediments and in oxygen-depleted zones of the ocean, fixed nitrogen is chemically reduced to N_2 by the microbial processes of denitrification and anaerobic ammonia oxidation. Understanding the relative balance of the fluxes of fixed nitrogen in the ocean requires the rate measurements of these competing processes to be accurate and well constrained.

On the source side of this nitrogen budget, the majority of published estimates of marine N_2-fixation rates[3] are based on a fairly straightforward protocol[4]: add a bubble of isotopically labelled nitrogen gas ($^{15}N_2$) to a sample of sea water; calculate the initial enrichment of $^{15}N_2$ in the sample using the ideal-gas law; incubate the sample for a specified period of time; and then measure the

*This article and the paper under discussion[2] were published online on 8 August 2012.

fraction of the isotopic tracer that is incorporated into cellular material in the sample. The net rate of N_2 fixation is then calculated by multiplying the fraction of biomass labelled per unit of time by the mass of nitrogen in particles suspended in the sample (which are assumed to be living material).

When these rates are extrapolated from localized study regions to the expanse of ocean thought to be habitable to diazotrophs, the resulting global N_2-fixation rates are generally found to be far less than those inferred by geochemical approaches[1]. This
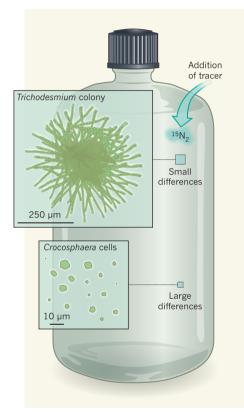
Figure 1 | Marine nitrogen fixation. A commonly used method for measuring the rate of nitrogen (N_2) fixation by diazotrophic microorganisms in the ocean is to add isotopically labelled nitrogen gas, $^{15}N_2$, to a sample of sea water, and then to measure the uptake of this nitrogen tracer by diazotrophs in the sample. The $^{15}N_2$ has classically been added as a bubble of gas[4], but Großkopf *et al.*[2] have compared the results of this traditional approach with those of a modified, more reliable version of the assay[10] — in which the tracer is added as a solution in filtered sea water — for measurements of the N_2-fixation rates of diazotrophs in the Atlantic Ocean. They found that the rates obtained using the modified method were, on average, greater than those obtained using the traditional approach. The differences between rates may in part track the composition of the diazotroph communities sampled — the discrepancies were larger when unicellular cyanobacteria, such as *Crocosphaera*, were in the majority, and smaller when the colony-forming genus *Trichodesmium* predominated.

potential discrepancy set scientists on a "relentless search"[5] for potentially missing sources of fixed nitrogen. These efforts paid off — they led to the discovery of a wider cast of nitrogen-fixing 'characters' than had been known before[6]; an expanded view of the ocean habitats occupied by such organisms[7]; the elucidation of previously unknown metabolic strategies for diazotrophic growth[8]; and descriptions of newly recognized symbioses between diazotrophs and other microorganisms[9]. Despite these revelations, the available N_2-fixation-rate measurements still indicated a deficit of oceanic fixed nitrogen. And so the question remained: if the nitrogen budget is in balance, then where have rate measurements of oceanic nitrogen fluxes gone wrong?

In 2010, a potential solution to this conundrum was reported[10] with the demonstration that when $^{15}N_2$ is added to a sample of sea water, it does not rapidly dissolve in the surrounding liquid. This was an unexpectedly simple solution — if a fraction of the added $^{15}N_2$ remains in the gas phase during fixation-rate measurements, instead of dissolving in sea water as was assumed, then one would underestimate N_2-fixation rates. To assess the effect of incomplete $^{15}N_2$ dissolution, laboratory experiments were performed in which the classic method was revised[10]: instead of adding a bubble of $^{15}N_2$ directly to a sample, the tracer was added to aliquots of filtered and degassed sea water, which were agitated until the bubble dissolved fully. The isotopic enrichment of the water was then verified by mass spectrometry[10]. By adding a known volume of this water to samples, rather than a bubble of $^{15}N_2$, the initial concentration of added tracer could be constrained, leading to more-accurate measurements. This methodological tweak resulted in significantly higher fixation rates being measured in the laboratory — a finding that, if applicable to *in situ* measurements, could substantially reduce the apparent imbalances previously observed for the marine fixed-nitrogen budget.

Großkopf and colleagues now report the first application of the modified 'dissolution' method to the analysis of natural populations of diazotrophs. During two research cruises in the Atlantic Ocean, the authors made 17 side-by-side comparisons of N_2-fixation rates obtained using the conventional bubble-addition approach[4] and the revised method[10]. Their major finding was that rates measured using the dissolution method were on average 1.7 times greater than those derived from the bubble method. Accordingly, the authors used this number in a first-pass calculation to adjust estimates of global nitrogen-fixation rates upward.

A secondary finding of this work[2] was that there was a "poor overall correlation between the N_2-fixation rates calculated using the two methods". The magnitude of

the difference between the methods varies, and in a few instances the methods actually generated similar values. Großkopf *et al.* discuss mechanisms that may explain this variation. Their explanations include the use of different time intervals between the injection of a $^{15}N_2$ bubble into a sample and the onset of the daily period of biological N_2 fixation, as well as the variable buoyancy of diazotrophs and hence their proximity to the $^{15}N_2$ bubble in the experiments. Intriguingly, both of these mechanisms seem to reflect the structure of diazotroph communities, to the extent that the identity of the dominant diazotroph along any cruise transect will affect the N_2-fixation rates measured using the bubble-addition technique (Fig. 1).

Großkopf and colleagues' findings will surely stimulate a spirited and necessary debate on how, and if, we can reinterpret the past decade or so of N_2-fixation measurements. In the past, the bubble-addition assay has been implemented inconsistently — different researchers used a range of incubation durations; experiments were initiated at different times of day; varying volumes of $^{15}N_2$ were injected; and experiments unavoidably targeted a spectrum of nitrogen-fixing communities. Given Großkopf and colleagues' results and this legacy of methodological inconsistency, some past measurements may need no revision, whereas others probably need large adjustments. Unfortunately, reliably identifying which measurements require revision may not be possible. The way forward is therefore clear: to understand the current state of the oceanic nitrogen cycle, researchers need to forge a path to a standardized methodology and to assess whether the database of historical N_2-fixation rates[3] can be adjusted, or whether field efforts will need to begin anew. ■

Angelicque E. White *is at the College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA.*
e-mail: awhite@coas.oregonstate.edu

1. Codispoti, L. A. *Biogeosci. Discuss.* **3,** 1203–1246 (2006).
2. Großkopf, T. *et al. Nature* **488,** 361–364 (2012).
3. Luo, Y.-W. *et al. Earth Syst. Sci. Data Discuss.* **5,** 47–106 (2012).
4. Montoya, J. P., Voss, M., Kahler, P. & Capone, D. G. *Appl. Environ. Microbiol.* **62,** 986–993 (1996).
5. Mahaffey, C., Michaels, A. F. & Capone, D. G. *Am. J. Sci.* **305,** 546–595 (2005).
6. Zehr, J. P., Mellon, M. T. & Zani, S. *Appl. Environ. Microbiol.* **64,** 3444–3450 (1998).
7. Deutsch, C., Sarmiento, J. L, Sigman, D. M., Gruber, N. & Dunne, J. P. *Nature* **445,** 163–167 (2007).
8. Zehr, J. P. *et al. Science* **322,** 1110–1112 (2008).
9. Foster, R. A. & O'Mullen, G. D. in *Nitrogen in the Marine Environment* 2nd edn (eds Capone, D. G. *et al.*) 1197–1218 (Elsevier, 2008).
10. Mohr, W., Großkopf, T., Wallace, D. W. R. & LaRoche, J. *PLoS ONE* **5,** e12583 (2010).

ARTICLE

Passenger deletions generate therapeutic vulnerabilities in cancer

Florian L. Muller[1,2,3]*, Simona Colla[1,2,3]*, Elisa Aquilanti[2]*, Veronica E. Manzo[2], Giannicola Genovese[1,2], Jaclyn Lee[2], Daniel Eisenson[2], Rujuta Narurkar[2], Pingna Deng[1,2], Luigi Nezi[1,2], Michelle A. Lee[2,4], Baoli Hu[1,2,5], Jian Hu[1,2,3], Ergun Sahin[2,3], Derrick Ong[1,2,3], Eliot Fletcher-Sananikone[1,2], Dennis Ho[2,3], Lawrence Kwong[1,2], Cameron Brennan[6], Y. Alan Wang[1,2,5], Lynda Chin[1,2,5] & Ronald A. DePinho[2,3,5,7]

Inactivation of tumour-suppressor genes by homozygous deletion is a prototypic event in the cancer genome, yet such deletions often encompass neighbouring genes. We propose that homozygous deletions in such passenger genes can expose cancer-specific therapeutic vulnerabilities when the collaterally deleted gene is a member of a functionally redundant family of genes carrying out an essential function. The glycolytic gene enolase 1 (ENO1) in the 1p36 locus is deleted in glioblastoma (GBM), which is tolerated by the expression of ENO2. Here we show that short-hairpin-RNA-mediated silencing of ENO2 selectively inhibits growth, survival and the tumorigenic potential of ENO1-deleted GBM cells, and that the enolase inhibitor phosphonoacetohydroxamate is selectively toxic to ENO1-deleted GBM cells relative to ENO1-intact GBM cells or normal astrocytes. The principle of collateral vulnerability should be applicable to other passenger-deleted genes encoding functionally redundant essential activities and provide an effective treatment strategy for cancers containing such genomic events.

Large-scale analysis of the cancer genome has provided an unprecedentedly detailed picture of the genetic anatomy of cancer[1], which has been, and continues to serve as, a blueprint for the development of molecular-targeted therapies. Targeted therapies directed against amplified or mutant-activated key driver oncoproteins have provided encouraging clinical progress[2], whereas exploiting loss-of-function mutations or gene deletions has received considerably less attention and has not been as successful thus far. Previous therapeutic work in the area of loss-of-function mutations and deletions has focused specifically on tumour-suppressor genes by strategies that include synthetic lethal approaches. One notable example of a synthetic lethal interaction is the response of BRCA1-mutant cancers to poly(ADP-ribose) polymerase inhibitors, although this interaction seems to be dependent on genetic context[3,4], and can be bypassed in late-stage tumours[3–5]. Most other synthetic lethal interactors of inactivated tumour suppressors seem to be less robust in eliciting cancer cell death[5], perhaps because such genes typically do not perform fundamental housekeeping functions.

Cancer genomes are characterized by numerous copy number amplifications and deletions, which target driver oncogenes and tumour-suppressor genes, respectively. Often, these genomic alterations are large regional events, affecting many other genes in addition to the intended target(s). The fact that such broad genomic alterations are not negatively selected against in cancer cells indicates that, on their own, the copy number alterations of these neighbouring passengers must not carry severely detrimental biological consequences. That said, it is conceivable that these passenger genomic events can create unintended (collateral) vulnerabilities unique to those cells; such as when a passenger being co-deleted is a member of a redundant multigene family serving an essential housekeeping function. A large body of genetic interaction studies in invertebrates

as well as mice indicates that many essential cellular housekeeping functions are carried out by several homologous genes that encode overlapping functions; this redundancy enables cell viability after loss of one homologue but causes lethality after the loss of several homologues[6–10] (Supplementary Fig. 1). In this conceptual framework, we proposed that the homozygous deletion of redundant essential housekeeping genes could create cancer-specific vulnerabilities (Supplementary Fig. 1a), in which pharmacological inactivation of the second, non-deleted homologue would result in the complete loss of activity in tumour cells carrying the deletion, without compromising the health of normal cells, in which both genes are intact and expressed (Supplementary Fig. 1b).

ENO1 is a redundant housekeeping gene deleted in GBM

By examining The Cancer Genome Atlas (TCGA) GBM data set for homozygous deletions targeting genes involved in essential cell activities[1], we identified various such candidates, including the ENO1 gene, which resides at the 1p36 tumour-suppressor locus (see Table 1 for a summary and Supplementary Table 1 for more detailed methodological support). Enolase, which is encoded by three homologous genes, is an essential enzyme that catalyses the second to last step of glycolysis, converting 2-phosphoglyceric acid into phosphoenolpyruvate[11]. In mammals, enolase activity is encoded by three genes: ENO1, which is ubiquitously expressed[12,13]; ENO2, which is expressed exclusively in neural tissues[12,14]; and ENO3, which is expressed in muscle tissues[15] (Supplementary Table 2). ENO1 is the major enolase isoform in GBM, accounting for 75–90% of cellular enolase activity[12]. Given the crucial importance of glycolysis for energy generation and anabolic processes in normal and especially tumour cells[16], GBM tumours homozygous null for ENO1 would be predicted to be highly sensitive to the inhibition of enolase 2, whereas

Table 1 | Homozygously deleted essential-redundant genes in GBM

Homozygously deleted gene	Chromosomal locus	Proximal tumour-suppressor gene	Target homologue	Pathway
ENO1	1p36.2	Various	*ENO2*	Glycolysis and gluconeogenesis
H6PD	1p36.2	Various	*G6PD*	Pentose phosphate shunt
KIF1B	1p36.2	Various	*KIF1A/C*	Chromosomal seggregation
NMNAT1	1p36.2	Various	*NMNAT2/3*	NAD⁺ biosynthesis
UBE4B	1p36.2	Various	*UBE4A*	Polyubiquitin-dependent degradation
ACO1	9p21.1	*INK/ARF*	*ACO2/3*	Regulation of iron metabolism/citric acid cycle
KLHL9	9p22	*INK/ARF*	*KLHL13*	Chromosomal segregation
PANK1	10q23.31	*PTEN*	*PANK3*	Acetyl-CoA biosynthesis
KIF20B	10q23.31	*PTEN*	*KIF20A*	Chromosomal segregation/cytokinesis

An evidence-based filtered list of genes homozygously deleted in GBM that are likely to execute an essential housekeeping function and have redundant (and potentially druggable) homologues. A more detailed methods description is provided in Supplementary Table 1.

normal neural tissues should not be affected because of the functional redundancy of enolase 1 (Fig. 1a, b). Correspondingly, *Eno2* knockout mice are viable and fertile, suggesting that pharmacological inhibition of enolase 2 is likely to be well tolerated at the organism level (Supplementary Table 2). Moreover, *Saccharomyces cerevisiae*, which possesses several enolase homologues, shows weak phenotypes with single mutants and incurs cell lethality only when all homologues are deleted[8–10]; whereas *Caenorhabditis elegans* and *Drosophila* possess only one gene encoding enolase activity, and its deletion is lethal[17,18].

The 1p36 locus, which contains several candidate tumour-suppressor genes, including *CHD5* and *CAMTA1* (refs 19, 20), sustains frequent deletion in GBM (Fig. 2a). The 1p36 locus is homozygously deleted in 1–5% of GBMs[1,21,22] (as well as oligodendrogliomas[23] and large-cell neuroendocrine lung tumours[24]), and *ENO1* is often included in the deletion. By examining the TCGA copy number aberrations (single nucleotide polymorphism (SNP) and array comparative genomic hybridization (aCGH) data)[1] and expression profiles, we identified 5 out of 359 GBM samples with homozygous deletion of *ENO1* and associated near-complete absence of its expression (Fig. 2b and Supplementary Fig. 2). We identified two GBM cell lines, D423-MG[22] and Gli56 (ref. 25), with homozygous deletions at the 1p36 locus spanning *ENO1*. A third GBM cell line,

D502-MG[22], also incurs homozygous deletion of many genes in this locus but not *ENO1*, thus serving as an excellent control (Fig. 2c). Western blot analysis confirmed the loss of ENO1 and the retention of ENO2 protein in D423-MG and Gli56 cells, whereas both proteins were present in D502-MG and in all other glioma and normal glial cell lines tested (Fig. 2d).

ENO2 knockdown inhibits growth of *ENO1*–deleted cells

We used the D502-MG (*ENO1* expressing) and D423-MG (*ENO1*-null) cell lines to assess the impact of short hairpin RNA (shRNA)-mediated knockdown of *ENO2* in an *ENO1* wild-type or null context. Two independent *ENO2* shRNAs (pLKO.1 vector) resulted in robust protein reduction and led to a profound inhibition of cell growth only in the context of *ENO1* genomic deletion (Supplementary Fig. 3). We obtained the same result using a further, independent *ENO2* shRNA (pGIPZ vector) (Supplementary Fig. 4a, b). Furthermore, shRNA ablation of *ENO2* in *ENO1*-null cells also resulted in decreased soft agar colony formation and blocked the *in vivo* tumorigenic potential of intracranially injected cells (Fig. 2e and Supplementary Fig. 4c). Finally, the selective toxicity of *ENO2* ablation to *ENO1*-null cells was demonstrated in an isogenic context using the doxycycline-inducible TRIPZ vector. When we used this doxycycline-inducible system in *ENO1* wild-type cell lines (U87, A1207 and LN319), two independent shRNAs reduced ENO2 protein levels by >70% (Fig. 3a) with no impact on ENO1 levels (data not shown). This *ENO2* ablation resulted in a profound inhibition of cell proliferation only in the *ENO1*-null D423-MG cell line (Fig. 3b). Furthermore, enforced expression of hairpin-resistant *ENO2* open reading frame fully reversed the deleterious effects of the *ENO2* shRNA (Supplementary Fig. 5), showing that the inhibitory effect of the hairpin was indeed specific to diminished ENO2 expression and was not an off-target effect. Finally, when ENO1 was ectopically re-expressed in D423-MG (*ENO1*-null) cell lines at levels similar to those observed in *ENO1* wild-type GBM lines, the deleterious effect of shRNA ablation of ENO2 was completely abrogated (Supplementary Fig. 6).

Enolase inhibition is toxic to *ENO1*–null cells

Next, we assessed the effect of pharmacological inhibition of enolase activity in *ENO1* wild-type and null cells. Previous studies have focused on the pharmacological inhibition of enolase, in particular for antiparasitic purposes[26,27], and many compounds have been characterized, most of which act as reaction-intermediate analogues (Supplementary Table 3). The most potent enolase inhibitor is phosphonoacetohydroxamate (PHAH)[27], which is thought to act as a transition-state analogue with an inhibitory constant of 15 pM on yeast enolase. Although PHAH has not been tested on human enolases, previous work demonstrated inhibitory effects on enolases from distantly related organisms[27,28], suggesting its potential use over a large phylogenetic distance. We find that PHAH was indeed capable of potent inhibition of enolase *in vitro* in native lysates of human GBM cell lines, with a half-maximum inhibitory concentration (IC_{50}) of around 20 nM (Fig. 4b and data not shown). We used PHAH in concentrations ranging from 0.625 μM to 50 μM and

Figure 1 | Homozygous deletions in *ENO1* sensitize tumours to molecular targeting of ENO2. a, *ENO1* is homozygously deleted in glioblastomas as part of the 1p36 locus. Loss of ENO1 is tolerable to the tumour because ENO2 is still expressed. **b**, A specific ENO2 inhibitor should completely eliminate enolase activity in *ENO1*-null tumour cells (hence blocking glycolysis and ATP synthesis), but leave genomically intact normal tissues unaffected because enolase activity is still present as ENO1 is still expressed.

Figure 2 | **Homozygous deletion of the 1p36 locus in GBM results in loss of *ENO1* expression in primary tumours and cell lines. a**, TCGA aCGH data show four primary GBMs with log2 copy number < −1, indicating homozygous deletion of the 1p36 locus. **b**, DNA copy number correlates with messenger RNA expression; expression is highest in tumours with n = 2 copies (wild type) and lowest in tumours with n = 0 copies (null) of *ENO1*. **c**, The D423-MG cell line was identified as homozygously deleted by SNP arrays from the Wellcome Trust Sanger Institute data set. **d**, The complete absence of ENO1 protein in D423-MG and Gli56 cells was confirmed by western blotting. **e**, shRNA targeting *ENO2* (shENO2-3) but not scramble (shNT) in D423-MG *ENO1*-null cells ablated intracranial tumorigenesis *in vivo* (n = 4 mice per group).

observed marked toxicity in *ENO1*-null cells (Fig. 4a, c and Supplementary Fig. 7) and minimal impact on the *ENO1* wild-type controls, which show at least ten times greater enolase activity relative to the *ENO1*-null cells (because ENO1 accounts for 90% of the total cellular enolase activity[12]; Fig. 4b). Although the IC_{50} of PHAH is similar for ENO1 and ENO2 in vitro (data not shown) the greater toxicity of the inhibitor to *ENO1*-null cells (Gli56 and D423-MG) derives from the fact that in these cells enolase activity is already 90% lower than in wild-type ENO1 cell lines, and consequently, a much lower dose is required to decrease total enolase activity below the toxicity threshold. Further data indicate a direct relationship between the levels of enolase activity and the sensitivity to PHAH across different cell lines and in the same cell line with different levels of enforced enolase expression. First, U343 and D502-MG cells, which have intermediate levels of enolase activity (and ENO1 protein expression; Fig. 2d) compared with the other cell lines, have intermediate levels of sensitivity to PHAH (Fig. 4), which in the case of U343 can be rescued by ectopic overexpression of *ENO1* or *ENO2* (data not shown). A systematic titration of PHAH in D423-MG cell lines with varying levels of enforced ENO1 or ENO2 expression, shows a direct relationship between the level of enolase expression/activity and the ensuing resistance to PHAH (Supplementary Fig. 7). PHAH toxicity was also abrogated in Gli56 *ENO1*-null cells by ectopic expression of physiological levels of *ENO1* or overexpression of *ENO2* (Supplementary Fig. 7). Regarding the mechanism of toxicity, cell cycle and apoptosis analysis demonstrated that PHAH treatment for 48 h induced a marked decrease of S-phase cells followed by a

marked increase of apoptosis in D423-MG but not in *ENO1* wild-type U373 cells (Supplementary Table 4). This effect was completely rescued by *ENO2* overexpression (data not shown). The fact that this growth inhibition and subsequent apoptosis is due to energy crisis is substantiated by a strong induction of phosphorylated 5′-AMP-activated protein kinase (AMPK) (at Thr 172)[29], which was observed in D423-MG but not in ENO1 wild-type cell lines (data not shown). It is tempting to speculate that this energy stress response exerts a protective effect, and thus the concomitant addition of an AMPK inhibitor together with PHAH could result in further toxicity. Finally, it is worth noting that *ENO1*-null cells do not show any greater sensitivity to other molecular-targeted therapies, such as a combination of receptor tyrosine kinase inhibitors[30] (lapatinib, sorafenib and PHA665752) (Supplementary Fig. 8) and rapamycin (data not shown) compared with wild-type *ENO1* cells. These data indicate that D423-MG cells are not broadly susceptible to other anticancer agents and that PHAH selectively targets *ENO1*-null GBM cells.

Discussion

In this study, we sought to determine the effect of collateral deletion of genes in tumour-suppressor loci that belong to redundant gene families playing cell-essential roles, and to assess whether extinction of remaining gene family members would create cancer-specific vulnerabilities. We provide genetic and pharmacological evidence that enolase 2 inhibition is lethal in cells with 1p36 homozygous deletion with collateral loss of *ENO1*, whereas *ENO1*-intact cells can rely on
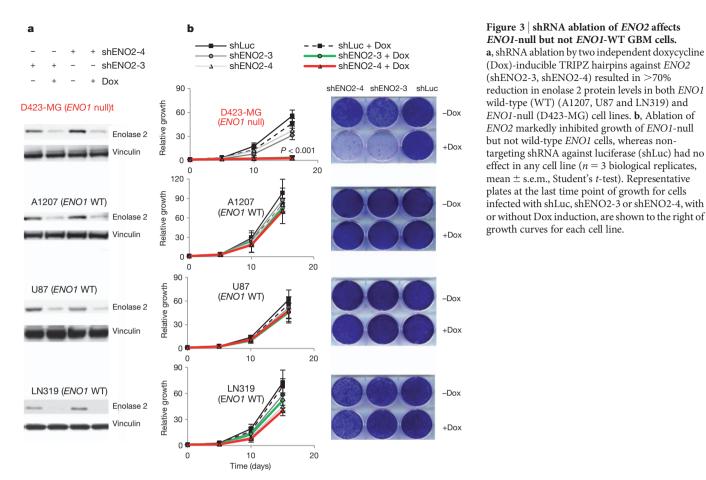
Figure 3 | shRNA ablation of *ENO2* affects *ENO1*-null but not *ENO1*-WT GBM cells.
a, shRNA ablation by two independent doxycycline (Dox)-inducible TRIPZ hairpins against *ENO2* (shENO2-3, shENO2-4) resulted in >70% reduction in enolase 2 protein levels in both *ENO1* wild-type (WT) (A1207, U87 and LN319) and *ENO1*-null (D423-MG) cell lines. **b**, Ablation of *ENO2* markedly inhibited growth of *ENO1*-null but not wild-type *ENO1* cells, whereas non-targeting shRNA against luciferase (shLuc) had no effect in any cell line (*n* = 3 biological replicates, mean ± s.e.m., Student's *t*-test). Representative plates at the last time point of growth for cells infected with shLuc, shENO2-3 or shENO2-4, with or without Dox induction, are shown to the right of growth curves for each cell line.
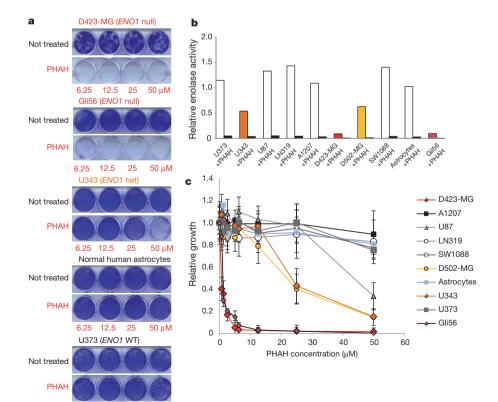
Figure 4 | Extreme sensitivity of *ENO1*-null cells to the pan-enolase inhibitor PHAH. a, D423-MG and Gli56 *ENO1*-null cell lines are highly sensitive to PHAH toxicity, whereas *ENO1* wild-type cell lines and normal astrocytes are not. Het, heterozygous deleted. **b**, The sensitivity of GBM lines to PHAH treatment correlated with their overall enolase activity. Pre-incubation of the lysates with 1 μM PHAH inhibited enolase enzymatic activity by >95% (average, *n* = 2 technical replicates). **c**, PHAH minimally affected the growth of *ENO1* wild-type GBM cells and normal astrocytes except at concentrations above 50 μM. Low PHAH concentrations stall the growth of *ENO1*-null cells, whereas cells with reduced *ENO1* expression (D502-MG and U343) showed intermediate sensitivity (*n* = 4 biological replicates, mean ± s.e.m.).

enolase 1 to undergo glycolysis and support survival. These findings are in agreement with genetic data from invertebrates[8–10]. Given that several homozygously deleted housekeeping genes can occur in the same deletion on 1p36 (for example, *H6PD*; Supplementary Table 1), it may be possible to increase the effectiveness and cancer cell-specific killing further by combining the inhibition of *ENO2* with that of another homologue of a simultaneously deleted housekeeping gene.

Attempts to therapeutically exploit general metabolic differences between normal and cancer cells, such as glucose addiction (the Warburg effect) and glutamine or serine addiction[16,31,32], remain areas of active preclinical investigation and clinical development. The approach described here is distinguished from these attempts in that it does not rely on any general aspect of cancer cell metabolism, but rather rests on genetically determined metabolic differences between normal and cancerous tissue to generate cancer cell-specific vulnerabilities. We propose that collateral vulnerability may be extended to other passenger homozygously deleted housekeeping genes in loci sustaining frequent deletion, such as 9p21 (*CDKN2A*) and 10q23 (*PTEN*), which contain members of functionally redundant housekeeping gene families (Table 1, Supplementary Table 1 and Supplementary Fig. 9). The strongest pharmacologically targetable candidates in this list are the homologues of the dual-function energy metabolism/iron regulator aconitase 1 (*ACO1*; 9p21) and the coenzyme A biosynthetic enzyme, pantothenate kinase 1 (*PANK1*; 10q23). Importantly, many of the compounds targeting the homologues of these passenger genes represent new molecular entities with respect to cancer treatment.

By one estimate, 11% of all protein-coding genes in the human genome are deleted in human cancers[33]. Thus, given the large number of homozygous deletions across many different cancer types spanning many hundreds of genes[33–36], the model described here for GBM should be applicable to the development of personalized treatments for many other cancer types.

METHODS SUMMARY

Cells were cultured using standard techniques in DMEM with 20% fetal bovine serum. shRNA experiments were conducted by lentiviral production through transient transfection of 293T cells followed by transduction in medium containing $4\,\mu g\,ml^{-1}$ polybrene and selection with $2\,\mu g\,ml^{-1}$ puromycin. shRNA expression was induced with $1\,\mu g\,ml^{-1}$ doxycycline and knockdown was tested by western blot. The shRNA-resistant ENO2 open reading frame was created by introducing silent mutations with the QuickChange site directed mutagenesis kit from Stratagene and then cloning into the pHAGE-CMV lentiviral vector. Cell proliferation experiments were performed using crystal violet staining, the CellTiter-Glo assay (Roche) and by measuring confluence using IncuCyte (Essen Bioscience). Orthotopic intracranial injections of D423-MG cells with and without ENO2 knockdown in severe combined immunodeficient (SCID) mice were performed as previously described[37]. Soft agar colony formation assay of these cells was performed using standard techniques by seeding 10^4 cells in 6-well plates. For the inhibitor studies, PHAH lithium salt was custom-synthesized by TCRS, following a previous protocol[28]. For the enolase activity assay, NADH oxidation was measured in a pyruvate kinase–lactate dehydrogenase coupled reaction as previously described[12]. For cell cycle studies, cells were incubated with or without PHAH for 48 h, stained with propidium iodide and sorted by flow cytometric analysis. For annexin V/7-AAD assays, cells were treated with or without PHAH for 96 h, stained with annexin V–phycoerythrin (PE) and 7-AAD, and evaluated for apoptosis by flow cytometry according to the manufacture's protocol (Biovision).

Full Methods and any associated references are available in the online version of the paper.

1. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455,** 1061–1068 (2008).
2. Druker, B. J. Translation of the Philadelphia chromosome into therapy for CML. *Blood* **112,** 4808–4817 (2008).
3. Guha, M. PARP inhibitors stumble in breast cancer. *Nature Biotechnol.* **29,** 373–374 (2011).
4. De Soto, J. A. & Deng, C. X. PARP-1 inhibitors: are they the long-sought genetically specific drugs for BRCA1/2-associated breast cancers? *Int. J. Med. Sci.* **3,** 117–123 (2006).
5. Weidle, U. H., Maisel, D. & Eick, D. Synthetic lethality-based targets for discovery of new cancer therapeutics. *Cancer Genomics Proteomics* **8,** 159–171 (2011).
6. Vavouri, T., Semple, J. I. & Lehner, B. Widespread conservation of genetic redundancy during a billion years of eukaryotic evolution. *Trends Genet.* **24,** 485–488 (2008).
7. Brookfield, J. F. Genetic redundancy. *Adv. Genet.* **36,** 137–155 (1997).
8. Costanzo, M. *et al.* The genetic landscape of a cell. *Science* **327,** 425–431 (2010).
9. DeLuna, A. *et al.* Exposing the fitness contribution of duplicated genes. *Nature Genet.* **40,** 676–681 (2008).
10. Deutscher, D., Meilijson, I., Kupiec, M. & Ruppin, E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nature Genet.* **38,** 993–998 (2006).
11. Poyner, R. R. & Reed, G. H. Structure of the bis divalent cation complex with phosphonoacetohydroxamate at the active site of enolase. *Biochemistry* **31,** 7166–7173 (1992).
12. Joseph, J., Cruz-Sanchez, F. F. & Carreras, J. Enolase activity and isoenzyme distribution in human brain regions and tumors. *J. Neurochem.* **66,** 2484–2490 (1996).
13. Stefanini, M. Chronic hemolytic anemia associated with erythrocyte enolase deficiency exacerbated by ingestion of nitrofurantoin. *Am. J. Clin. Pathol.* **58,** 408–414 (1972).
14. Kobayakawa, K. *et al.* Innate versus learned odour processing in the mouse olfactory bulb. *Nature* **450,** 503–508 (2007).
15. Comi, G. P. *et al.* β-enolase deficiency, a new metabolic myopathy of distal glycolysis. *Ann. Neurol.* **50,** 202–207 (2001).
16. Wise, D. R. & Thompson, C. B. Glutamine addiction: a new therapeutic target in cancer. *Trends Biochem. Sci.* **35,** 427–433 (2010).
17. Buszczak, M. *et al.* The carnegie protein trap library: a versatile tool for *Drosophila* developmental studies. *Genetics* **175,** 1505–1531 (2007).
18. Sonnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434,** 462–469 (2005).
19. Henrich, K. O. *et al.* CAMTA1, a 1p36 tumor suppressor candidate, inhibits growth and activates differentiation programs in neuroblastoma cells. *Cancer Res.* **71,** 3142–3151 (2011).
20. Bagchi, A. & Mills, A. A. The quest for the 1p36 tumor suppressor. *Cancer Res.* **68,** 2551–2556 (2008).
21. Yin, D. *et al.* High-resolution genomic copy number profiling of glioblastoma multiforme by single nucleotide polymorphism DNA microarray. *Mol. Cancer Res.* **7,** 665–677 (2009).
22. Duncan, C. G. *et al.* Integrated genomic analyses identify ERRFI1 and TACC3 as glioblastoma-targeted genes. *Oncotarget* **1,** 265–277 (2010).
23. Kotliarov, Y. *et al.* High-resolution global genomic survey of 178 gliomas reveals novel regions of copy number alteration and allelic imbalances. *Cancer Res.* **66,** 9428–9436 (2006).
24. Peng, W. X. *et al.* Array-based comparative genomic hybridization analysis of high-grade neuroendocrine tumors of the lung. *Cancer Sci.* **96,** 661–667 (2007).
25. Mueller, W. *et al.* Downregulation of RUNX3 and TES by hypermethylation in glioblastoma. *Oncogene* **26,** 583–593 (2005).
26. Guha-Chowdhury, N., Clark, A. G. & Sissons, C. H. Inhibition of purified enolases from oral bacteria by fluoride. *Oral Microbiol. Immunol.* **12,** 91–97 (1997).
27. de A. S. & Navarro, M. V. *et al.* Structural flexibility in *Trypanosoma brucei* enolase revealed by X-ray crystallography and molecular dynamics. *FEBS J.* **274,** 5077–5089 (2007).
28. Anderson, V. E., Weiss, P. M. & Cleland, W. W. Reaction intermediate analogues for enolase. *Biochemistry* **23,** 2779–2786 (1984).
29. Jing, M. & Ismail-Beigi, F. Critical role of 5′-AMP-activated protein kinase in the stimulation of glucose transport in response to inhibition of oxidative phosphorylation. *Am. J. Physiol. Cell Physiol.* **292,** C477–C487 (2007).
30. Stommel, J. M. *et al.* Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. *Science* **318,** 287–290 (2007).
31. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476,** 346–350 (2011).
32. Raj, L. *et al.* Selective killing of cancer cells by a small molecule targeting the stress response to ROS. *Nature* **475,** 231–234 (2011).
33. Bignell, G. R. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463,** 893–898 (2010).
34. Taylor, B. S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18,** 11–22 (2010).
35. Cox, C. *et al.* A survey of homozygous deletions in human cancer genomes. *Proc. Natl Acad. Sci. USA* **102,** 4542–4547 (2005).
36. Tonon, G. *et al.* High-resolution genomic profiles of human lung cancer. *Proc. Natl Acad. Sci. USA* **102,** 9625–9630 (2005).
37. Zheng, H. *et al.* p53 and Pten control neural and glioma stem/progenitor cell renewal and differentiation. *Nature* **455,** 1129–1133 (2008).

METHODS

Cell culture. The cell lines D423-MG (1p36 homozygously deleted, including *ENO1*) and D502-MG (1p36 homozygously deleted, excluding *ENO1*) were provided by D. Bigner[22]. (D423 and D502 are referred to as H423 and H502, respectively, in ref. 22, but as D423-MG and D502-MG in the Wellcome Trust Sanger Institute database (http://www.sanger.ac.uk), the nomenclature we adopt here). Gli56 was obtained from D. N. Louis as described in ref. 25. The deletion in D423-MG spans the *CAMTA1*, *VAMP3*, *PER3*, *UTS2*, *TNFRSF9*, *PARK7*, *ERRFI1*, *SLC45A1*, *RERE*, *ENO1*, *CA6*, *SLC2A5*, *GPR157*, *MIR34A*, *H6PD*, *SPSB1* and *SLC25A33* genes, whereas the deletion in Gli56 spans the *UTS2*, *TNFRSF9*, *PARK7*, *ERRFI1*, *SLC45A1*, *RERE*, *ENO1*, *CA6*, *SLC2A5*, *GPR157*, *MIR34A*, *H6PD*, *SPSB1*, *SLC25A33*, *TMEM201*, *C1orf200*, *PIK3CD*, *CLSTN1*, *CTNNBIP1*, *LZIC*, *NMNAT1*, *RBP7* and *UBE4B* loci. Cells were cultured in DMEM with 20% fetal bovine serum (FBS). For comparison, the cell lines U87, LN319, SW1088, U343, U373 and A1207 were grown under the same conditions. Normal human astrocytes were obtained from ScienCell.

shRNA knockdown of *ENO2* expression. We screened 22 hairpins targeting *ENO2* and found four independent ones that reduced protein levels by <50%. Two of these hairpins were in the pLKO.1 vector (shENO2-1 and shENO2-2), and the remaining two in the Expression Arrest GIPZ (shENO2-3) and TRIPZ (shENO2-4) shRNAmir vectors (Open Biosystems). The *ENO2* shRNA sequences are as follows: shENO2-1: 5′-CAAGGGAGTCATCAAGGACAA-3′; NCBI accession NM_001975. shENO2-2: 5′-CGCCTGGCTAATAAGGC TTTA-3′; NM_001975. shENO2-3: 5′-CGGCCTTCAACGTGATCAA-3′; NM_001975. shENO2-4: 5′-GGGACTGAGAACAAATCCA-3′; NM_001975.

The hairpin in the GIPZ vector was cloned into the TRIPZ vector using a protocol provided by the manufacturer. The TRIPZ vector is a doxycycline-inducible system with a red fluorescent protein reporter that is expressed only after doxycycline induction. Recombinant lentiviral particles were produced by transient transfection of 293T cells following a standard protocol. In brief, 72 µg of the shRNA plasmid, 54 µg of delta 8.9 plasmid and 18 µg of the VSVG plasmid were transfected using FuGene (Roche) into 293T cells plated in 245-mm^2 dishes. Viral supernatant was collected 72 h after transfection, concentrated by ultracentrifugation at 90,000*g*, and resuspended in cell growth medium. For transduction, viral solutions were added to cell culture medium containing 4 µg ml^{-1} polybrene; 48 h after infection, cells were selected using 2 µg ml^{-1} puromycin and tested for *ENO2* knockdown by western blotting.

Proliferation assays and anchorage-independent growth. Cell growth of shRNA- or PHAH-treated cell lines was assayed either through crystal violet staining or using the Promega CellTiter-Glo proliferation kit (Roche) or alternatively, *in vivo*, by measuring confluence with the IncuCyte (Essen BioScience). Growth curves using the IncuCyte were generated by imaging every 2 h with quadruplicate replicates. For crystal violet assays, 10^4 cells were seeded in a 6-well plate for each time point. At the indicated time point, cells were fixed with 10% formalin and stained with crystal violet solution for 1 h. Dye extraction was performed using 10% acetic acid solution, and absorbance was read at 590 nm. CellTiter-Glo experiments were performed according to the manufacturer's instructions; 10^3 cells per well were plated in a 96-well plate for each time point, and luminescence readings were taken every 24 h. All experiments were performed in triplicate. Soft agar (anchorage-independent) growth was monitored in 6-well plates seeded with 10^4 cells of the indicated genotype. The medium contained DMEM with 10% FBS; the top agar contained 0.4% low melting agarose, whereas the bottom agar contained 1% low melting agarose. Growth was monitored by fluorescence (green fluorescent protein, GFP) and after 28 days colonies were stained with iodonitrotetrazolium chloride (Sigma-Aldrich) and counted.

Orthotopic brain tumour formation. The *in vivo* tumorigenic potential of D423-MG cells transduced with non-targeting hairpin or shENO2-3 delivered

through pGIPZ was determined as previously described[37]. SCID mice (Charles River) under deep anaesthesia were placed into a stereotactic apparatus equipped with a *z*-axis (Stoelting). Then, 3×10^5 cells were injected intracranially into the right caudate nucleus, 3 mm below the surface of the brain, using a 10-µl Hamilton syringe. The animals were followed daily for the development of neurological deficits. All mice experiments were performed with the approval of the Harvard Cancer Center and Dana-Farber Cancer Institute Institutional Animal Care and Use Committee.

Enolase activity assay. Enolase activity was measured by NADH oxidation in a pyruvate kinase–lactate dehydrogenase coupled assay as previously described[12]. In brief, cells were lysed in 20 mM Tris-HCl, 1 mM EDTA and 1 mM β-mercaptoethanol (pH 7.4), and homogenized using a Polytron homogenizer three times for a period of 10 s followed by sonication. Enolase activity was recorded by measuring oxidation of NADH either spectrophotometrically by absorbance at 340 nm or fluorescently by excitation at 340 nm and emission at 460 nm.

Western blotting. After two washes with PBS, cells were incubated in RIPA buffer for 15 min with gentle shaking. Lysates were then collected, sonicated and centrifuged at 20,000*g* for 10 min at 4 °C. SDS–PAGE and western blotting were performed as described previously[37]. The following antibodies were used: enolase 1 (3810), enolase 2 (9536), GAPDH (3683) and phosphor-AMPK Thr 172 (2535) from Cell Signaling Technologies, and vinculin from Sigma-Aldrich.

Inhibitor studies. PHAH lithium salt was custom-synthesized by TCRS, following the protocol described previously[28]. Structure and purity were verified by NMR. PHAH was dissolved in PBS at 50 mM stock and stored frozen at −80 °C until use. Given the instability of the compound, the medium was replaced every 5 days and fresh inhibitor added with fresh medium. Rapamycin, sorafenib, lapatinib were obtained from LC Laboratories and PHA665752 from Tocris Bioscience, respectively.

Ectopic expression of *ENO1*, *ENO2* and shRNA-resistant *ENO2*. Rescue of the phenotypic effects of knocking down *ENO2* in the cell line D423-MG was performed by overexpressing an shRNA-resistant form of *ENO2*. In brief, six silent mutations were introduced into the *ENO2* coding region targeted by shENO2-4, using the QuikChange site-directed mutagenesis kit (Stratagene). The shRNA-resistant *ENO2* coding region was cloned into the pHAGE-CMV lentiviral vector (a gift from D. N. Kotton) and overexpressed in the D423-MG cell line carrying shENO2-4, in the presence or absence of doxycycline. As a control, the same cell line was infected with a lentiviral vector carrying the GFP gene. For the ectopic re-expression of *ENO1* or *ENO2*, sequenced verified cDNA clones were gateway cloned into the pHAGE-CMV lentiviral vector and lentivirally transduced into glioma cell lines as described earlier.

Cell cycle analysis. The D423-MG and U373 cell lines were treated for 48 h in the presence or absence of PHAH (25 µM) and fixed in 75% ethanol at −20 °C overnight. The next day, the cells were washed with cold PBS, treated with 100 µg of RNase A (Qiagen), and stained with 50 µg of propidium iodide (Roche). Flow cytometric acquisition was performed using a three-colour FACScan flow cytometer and CellQuest software (Becton Dickinson). For each sample, 10^4 events were gated. Data analysis was performed using ModFit LT (Verity Software House).

Annexin V/7-AAD assay for apoptosis. The D423-MG and U373 cell lines were treated for 96 h in the presence or absence of PHAH (25 µM). For Annexin V/7-AAD assay cells were stained with annexin V–PE and 7-AAD, and evaluated for apoptosis by flow cytometry according to the manufacturer's protocol (Biovision). The apoptotic cells were determined using a Becton Dickinson FACScan cytometer. Both early apoptotic (annexin V-positive and 7-AAD-negative) and late apoptotic (annexin V-positive and 7-AAD-positive) cells were included in cell death determinations.

ARTICLE

Division and subtraction by distinct cortical inhibitory networks *in vivo*

Nathan R. Wilson[1]*, Caroline A. Runyan[1]*, Forea L. Wang[1] & Mriganka Sur[1]

Brain circuits process information through specialized neuronal subclasses interacting within a network. Revealing their interplay requires activating specific cells while monitoring others in a functioning circuit. Here we use a new platform for two-way light-based circuit interrogation in visual cortex *in vivo* to show the computational implications of modulating different subclasses of inhibitory neurons during sensory processing. We find that soma-targeting, parvalbumin-expressing (PV) neurons principally divide responses but preserve stimulus selectivity, whereas dendrite-targeting, somatostatin-expressing (SOM) neurons principally subtract from excitatory responses and sharpen selectivity. Visualized *in vivo* cell-attached recordings show that division by PV neurons alters response gain, whereas subtraction by SOM neurons shifts response levels. Finally, stimulating identified neurons while scanning many target cells reveals that single PV and SOM neurons functionally impact only specific subsets of neurons in their projection fields. These findings provide direct evidence that inhibitory neuronal subclasses have distinct and complementary roles in cortical computations.

Inhibition has fundamental and diverse roles in brain function, and is delivered by specialized cell types with distinct intrinsic properties and connectivity patterns[1–3]. This heterogeneity in cellular form and function suggests that different inhibitory subtypes may actually underpin distinct computational functions and even hold specific relevance to neurological disorders[4] based on their unique morphologies and functional positions within the network. Previous pharmacological or intracellular studies in primary visual cortex (V1), which necessarily considered inhibition as a single entity, have produced diverse findings on the role of inhibition. On the one hand, inhibition has been proposed to sharpen neuronal responses by removing weak inputs[5–7], though there have been conflicting reports on whether inhibition predominantly targets non-preferred responses or preferred ones[8–10]. On the other hand, inhibition has been posited to control response gain, a network mechanism by which cortical networks rapidly 'divide' or scale their dynamic range of responses[11]. This mechanism has been proposed as fundamental to processing across many brain systems, from primary sensory computations[12] to attention[13], multisensory integration[14], and value estimation[15].

Here we show that inhibition in the cerebral cortex can have either of these functions, depending on its cellular source. We propose that two key inhibitory neuron subclasses, soma-targeting PV neurons and dendrite-targeting SOM neurons, which together comprise a substantial proportion of cortical inhibitory neurons in mice[16,17], drive different kinds of inhibition. We combined optogenetic activation of individual or populations of PV or SOM neurons with monitoring the effects in target cells using high-speed imaging of functional responses[18] as well as cell-attached electrophysiological recordings[19]. These methods complement both static wiring diagrams[20] and wiring patterns examined in tissue slices[21–24] by revealing targeting specificity and functional consequences of inhibitory neuron activation in intact circuits processing visual information.

Optical dissection of network interactions

To measure the effects of distinct cell classes within a functioning network, we built a custom system combining optogenetic stimulation with *in vivo* two-photon imaging in the mammalian brain (Fig. 1a and Supplementary Fig. 1). Our imaging system (Supplementary Movie 1) sampled calcium responses from neurons loaded with a fluorescent reporter using a scan path customized for each image[25], at high speed but also high dwell times within neurons, yielding highly repeatable measurements of orientation-selective responses and clear tuning curves (Fig. 1a–f, Supplementary Figs 2 and 3, and Supplementary Movie 2).

To optically activate PV or SOM neurons, in parallel experiments, we used Cre/*loxP* recombination to express channelrhodopsin-2 (ChR2) in PV or SOM neurons in the mouse visual cortex (Supplementary Fig. 4). This led to highly specific and reliable on-demand activation of infected neurons in visual cortex that was verified both in slices and during visual stimulation *in vivo* (Fig. 1g and Supplementary Fig. 4). PV or SOM neurons were photo-activated for a 1-s interval at the onset of visual stimulation (Fig. 1h), enabling us to compare the control visual responses of neighbouring, non-infected neurons during episodically presented drifting oriented gratings (Fig. 1i) to responses in interleaved trials in which the cells were inhibited through PV or SOM activation (Fig. 1j and Supplementary Fig. 5). Concurrent calcium imaging and optogenetic stimulation enabled us to quantify interneuron suppression of neighbouring cells across the network (Fig. 1k).

Distinct functions of inhibitory cell classes

Using this system, we activated PV or SOM cells while recording the visual responses of non-infected cells to oriented drifting gratings (Supplementary Movie 3, Fig. 2a–d and Supplementary Fig. 5). Control responses of target cells were similar in PV and SOM experiments (Supplementary Fig. 6), and suppression by ChR2 was calibrated to a similar moderate range for all experiments (Supplementary Fig. 6i). Interestingly, PV activation caused a larger suppression when control responses were higher (Fig. 2e), whereas SOM activation caused a relatively uniform suppression of the full response profile (Fig. 2f), particularly when the control responses or baseline levels were high relative to the suppression thereby avoiding a 'floor effect'. Indeed,

[1]Department of Brain and Cognitive Sciences, Picower Institute for Learning and Memory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139 USA.
*These authors contributed equally to this work.
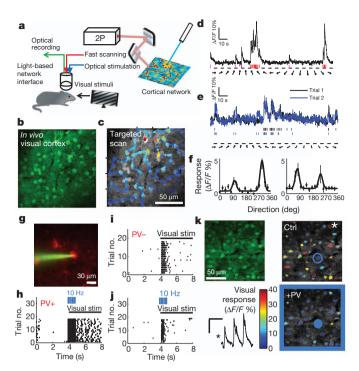
Figure 1 | All-optical network dissection of cortical subclasses during visual computations. a, Directed light, for optical recording and stimulation, was targeted to the V1 of an animal being shown visual stimuli. **b,** A traditional raster scan located cells in the network that were bulk-loaded with the calcium indicator dye OGB. **c,** Cells that were automatically identified were then imaged at high speed using targeted two-photon scanning along an arbitrary scan path. **d,** This enabled detection of robust cellular activity in response to episodically presented oriented drifting gratings, and this activity could be analysed as primary signals or deconvolved to estimate action potentials (red lines). **e,** Evoked optical traces were highly consistent over repeated presentations of visual stimuli. **f,** The clear, repeatable responses enabled the resolution of well-defined responses (dots), fit by dual Gaussian curves (lines). Data are shown as mean ± s.e.m. **g,** Image of an mCherry-ChR2+ PV+ (channelrhodopsin-2- and PV-positive) cell targeted *in vivo* for cell-attached recording. **h,** Evoked action potentials from an mCherry-ChR2+ PV+ cell. PV+ neurons were activated at the onset (3.9 s) of 4 s of visual stimulation, through 10-Hz stimulation of the PV+ neuron. **i, j,** Recording of the visual response of a PV− neuron in the control condition (**i**) and with PV activation (**j**). **k,** Left, cell population loaded with OGB dye (top), and responses of a cell marked by an asterisk in the network shown in the top right panel (bottom). Right, cells colour-coded by the magnitude of their visual response in the control condition (top) and when PV+ neurons were activated (bottom; colour bar, ΔF/F %). A ChR2+ PV neuron in the network is circled. +PV, optical PV activation; Ctrl, control; PV−, PV-negative; PV+, PV-positive.

the slope of the relationship between relative suppression and control response was significantly greater when PV cells were activated (Fig. 2g, i) than when SOM cells were activated (Fig. 2h, i; PV, $n = 150$ cells; SOM, $n = 77$ cells; $P < 0.001$, Kolmogorov–Smirnov test). Comparing the suppression at different response strengths for pooled cells further showed the asymmetric relationship between response and suppression when PV cells were activated (Fig. 2j; responses at 40, 60 and 80% of the maximum were less suppressed than those at maximum response strength; $P < 0.05$–0.001 for all pair-wise comparisons with 100%). In contrast, SOM suppression affected weak and strong responses similarly (Fig. 2k; responses at 40, 60 and 80% compared to responses at 100%; $P > 0.2$). Thus, suppression by SOM cells is relatively uniform across responses of different strengths, whereas suppression by PV cells is non-uniform and proportional to the response level of the target neuron.

Proportionate suppression would 'scale' responses, reducing the tuning curve's peak more strongly than its spontaneous response level or baseline, whereas uniform suppression would 'shift' the entire

tuning curve downwards, including the baseline (Supplementary Fig. 7). Indeed, averaging the tuning curves of cells recorded during PV-cell ChR2 activation revealed a scaled down version of the control tuning curve (Fig. 2l), whereas activating SOM-cell ChR2 yielded a more uniform downwards shift of the control curve (Fig. 2m).

The presence of a response 'floor' or threshold influences how different forms of suppression impact the orientation tuning curve (Supplementary Fig. 7). Examining cells in which post-ChR2 responses were largely above the 'floor', to accurately quantify the full distribution of suppression across the whole curve (Supplementary Fig. 7a), showed that PV and SOM activation both reduced the baseline responses at non-preferred orientations (Fig. 2n; +PV (PV neurons activated optically), 11 ± 4% decrease in baseline, $P < 0.05$, $n = 52$ cells; versus +SOM, 19 ± 6%, $P < 0.001$, $n = 25$ cells). PV but not SOM activation decreased the peak-baseline amplitude of cells' tuning functions (Fig. 2o; +PV, 28.5 ± 4.5% reduction, $P < 0.001$; +SOM, 19.8 ± 11.2%, $P = 0.10$), consistent with a larger reduction at the peak relative to the baseline. PV activation did not affect the orientation selectivity index (OSI) of target cells (Fig. 2p; control, 0.29 ± 0.01; versus +PV, 0.29 ± 0.01; $P = 0.71$; see also Supplementary Fig. 8), whereas SOM activation increased the OSI (Fig. 2p; control, 0.26 ± 0.01; versus +SOM, 0.30 ± 0.01; $P < 0.01$). Similarly, PV activation did not affect the tuning width of target cells (Fig. 2q; control, half-width at half-height, 38.7 ± 3.0 degrees; versus +PV, 35.2 ± 3.0 deg; $P = 0.31$), whereas SOM activation narrowed the tuning width (Fig. 2q; control, 45.7 ± 4.1 deg; versus +SOM, 37.5 ± 4.3 deg; $P < 0.01$). The effects on the direction selectivity index (DSI) were similar in trend (Fig. 2r; PV: 0.33 ± 0.03 control DSI versus 0.36 ± 0.03 +PV DSI, $P = 0.44$; SOM: 0.28 ± 0.04 control DSI versus 0.34 ± 0.05 +SOM DSI, $P = 0.10$). Thus, the relatively uniform suppression by SOM cells leads to a sharpening in response selectivity of target neurons, whereas the non-uniform but proportional suppression by PV cells reduces response magnitude but does not change response selectivity.

PV and SOM effects measured electrophysiologically

We examined further the different effects of PV and SOM activation using electrophysiological cell-attached recordings *in vivo* (Fig. 3a). Putative pyramidal neurons (Fig. 3b) were identified by their regular spiking properties (Fig. 3c); the peak:trough ratio of individual spikes was larger for all recorded cells than in identified fast-spiking PV-positive neurons (2.82 ± 0.20 for all recorded cells, $n = 21$; 1.41 ± 0.18 in PV-positive neurons, $n = 53$; $P < 0.001$). Moderate levels of PV- and SOM-mediated suppression (Supplementary Fig. 6i) had clearly different effects on spike responses of target cells and resultant orientation tuning curves (Fig. 3d, e), even in cells with very different response levels (Fig. 3f, g). PV suppression depended on the level of control response, whereas SOM suppression shifted tuning curves downwards more uniformly (Fig. 3h–k; PV, $n = 21$ cells, SOM, $n = 17$ cells). The changes in average tuning curves (Fig. 3l, m), and their parameters (Fig. 3n–r; Supplementary Fig. 8), showed that SOM but not PV suppression sharpened response selectivity, consistent with a model in which PV activation leads to a division of target cell responses but in which SOM activation leads to a subtraction (Supplementary Fig. 9).

Differential inhibitory impact on target-cell gain

The divisive impact of PV activation suggests that PV neurons implement dynamic response gain control in cortex, which has been previously attributed to intracortical inhibition[26,27]. We carried out cell-attached recordings and examined the effects of PV and SOM activation on a canonical measure of response gain, the modulation of responses with increasing contrast (Fig. 4a, b). PV activation (Fig. 4c) led to contrast response curves with reduced gain (slope), whereas SOM activation (Fig. 4d) decreased responses relatively uniformly (with a floor effect at low response levels). The PV activation curves
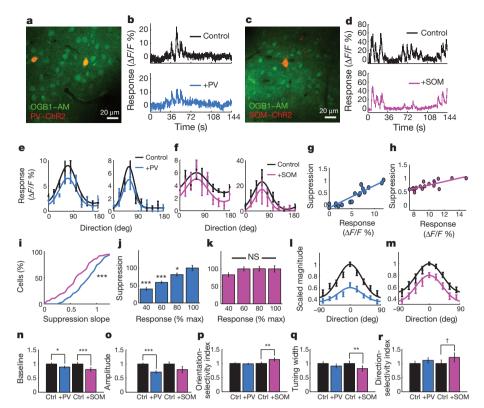
Figure 2 | Impact of PV- and SOM-driven inhibition on the tuning of neuronal responses. **a**, An imaging site showing neurons loaded with calcium indicator (OGB1-AM, green) and two PV+ neurons expressing mCherry-ChR2 (PV-ChR2, red) in visual cortex *in vivo*. **b**, Optical responses to visual stimuli, either without (Control) or with (+PV) simultaneous optical PV activation in interleaved trials, recorded during episodically presented oriented drifting gratings (see Fig. 1d). The photo-artefact from ChR2 stimulation has been removed from these trials. **c**, **d**, Same as (**a** and **b**), but from an experiment with mCherry-ChR2 expression in SOM neurons. **e**, **f**, Control tuning curves (black) were suppressed with ChR2 activation of PV neurons (blue) or SOM

neurons (pink) in four example cells. **g**, **h**, The normalized suppression plotted as a function of control response strength in two example cells (blue, PV; pink, SOM). **i**, Cumulative density functions of the distributions of suppression versus response slopes for all cells suppressed by PV (blue) and SOM (pink) activation. **j**, **k**, Population data showing the amount of PV- and SOM-mediated suppression at different response levels. **l**, **m**, Average tuning curves showing control responses and effects of PV (blue) or SOM (pink) activation. **n–r**, Effects of PV and SOM activation on tuning, including baseline, peak-baseline amplitude, OSI, tuning width and DSI. NS, not significant. †$P < 0.10$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Data are shown as mean ± s.e.m.

were better fit by a divisive scaling model than a subtractive model, whereas SOM activation curves were better fit by the subtractive model (Supplementary Fig. 10). PV- but not SOM-mediated suppression was dependent on the response level, as shown by both suppression-response slopes (Fig. 4e, f; PV, 0.11 ± 0.03, $n = 17$ cells; versus SOM, -0.008 ± 0.02, $n = 16$ cells; $P < 0.01$; Supplementary Fig. 11a, b) and suppression-response strength comparisons (PV, $P < 0.05$–0.001 comparing suppression at 40%, 60% and 80% response to that at 100% (Fig. 4g); SOM, $P > 0.2$ comparing suppression at 40%, 60% and 80% response to that at 100% (Fig. 4h)). PV activation scaled response magnitude (Fig. 4i; +PV: R_{max} $64.1 \pm 3.5\%$ of control; $n = 17$ cells; $P < 0.001$) without affecting half-saturation contrast (Fig. 4j; +PV: C_{50} $95.0 \pm 10\%$ of control; $P = 0.65$). SOM activation significantly reduced R_{max} (Fig. 4i; +SOM: $73.6 \pm 5.3\%$ of control; $n = 16$ cells; $P < 0.001$), but unlike PV, also significantly increased C_{50} (Fig. 4j; +SOM: $141 \pm 17.7\%$ of control; $P < 0.05$), with no effect on the response gain ($P = 0.23$ comparing slopes at C_{50} before and after SOM activation; versus $P < 0.01$ comparing slopes before and after PV activation; see also Supplementary Fig. 11). Thus, PV but not SOM activation contributes directly to controlling the gain of target-cell responses.

Single-cell circuit maps of network connections *in vivo*

The functional roles of inhibitory neurons are manifested through the spatial distribution and functional targeting of subclasses of inhibition onto cortical cells in the local network. To define the output connections of single inhibitory neurons, we developed a system to focally

stimulate an individual neuron while simultaneously imaging responses from large numbers of cells to assess their functional coupling (Fig. 5a). The ChR2-stimulating 473-nm beam was narrowed to a small effective radius and focused on sparsely distributed ChR2-positive cells (Fig. 5b and Supplementary Figs 12–16). Thus, we could optically activate chosen PV or SOM neurons *in vivo* while concurrently sampling population responses with targeted imaging (Fig. 5c and Supplementary Movie 4).

Mapping response modulation across a network while controlling a PV ChR2 cell or a SOM ChR2 cell (Fig. 5d, h), we found that visual responses of some neighbouring neurons were significantly suppressed while other cells were unaffected (Fig. 5e, i), resulting in maps of the functional suppression triggered during focal PV or SOM activation (Fig. 5f, j). Similar non-uniform, heterogeneous maps of functional PV and SOM cell connectivity were obtained from every imaged animal (Supplementary Fig. 17; PV, $n = 4$ networks; SOM, $n = 5$ networks). Focal PV activation resulted in the significant suppression of $43.1 \pm 2.1\%$ of neurons within the field of view, whereas focal SOM activation suppressed $16.2 \pm 2.9\%$ of neurons. Electrical stimulation of a single cell through whole-cell patch recording *in vivo* yielded a similar suppression map (Supplementary Fig. 18). Single PV or SOM neuron activation rarely triggered observable dis-inhibition (Supplementary Fig. 19). The nature of suppression was very similar to that observed with full-field activation: the amount of suppression by PV cells depended on the strength of the control response, whereas focal SOM stimulation resulted in more uniform suppression (Fig. 5h, m; see also Supplementary Fig. 20). Thus, the effects of SOM- and

PV-mediated inhibition are distinct, whether they are evoked by populations or single SOM or PV neurons. Furthermore, these maps of affected neurons show remarkable diversity in the functional suppression exerted by specific PV or SOM neurons within their local neighbourhoods.
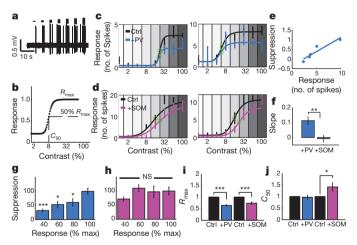
Figure 4 | Modulation of response gain by PV and SOM cells during targeted cell-attached recordings. a, Example responses from a neuron stimulated with drifting gratings of increasing contrast, at the cell's preferred orientation. **b**, Naka–Rushton curve describing the increase in responses with increasing contrast. **c, d**, Four example cells whose control response–contrast curves (black) were suppressed by PV (blue) or SOM (pink) activation. C_{50} values are marked (green circles). **e**, Relationship between suppression and response when a PV cell was activated (slope = 0.15). **f**, The mean suppression versus response slope was larger when PV cells were activated than when SOM cells were activated. **g, h**, Population data for the amount of PV- and SOM-mediated suppression observed at different response levels. **i, j**, Effects of PV and SOM activation on response–contrast function parameters R_{max} and C_{50} (see **b**). *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; NS, not significant. Data are shown as mean ± s.e.m.

Functional connectivity of inhibitory networks

An important question is whether there is an underlying logic through which an inhibitory cell makes functional connections with its target cells. We tested whether, *in vivo*, functional suppression by single PV or SOM neurons can be predicted by spatial[22,23] or functional[28,29] relationships. The distance to potential targets did not predict whether a neuron was significantly ($P < 0.05$) suppressed by PV or SOM neurons within 100 μm (Supplementary Fig. 21a), and there was no significant relationship between suppression strength and distance for the population of either PV- or SOM-stimulated networks (Fig. 6a; PV, $P = 0.80$, $R = -0.03$, 4 networks; SOM, $P = 0.1$, $R = -0.28$, 5 networks). Furthermore, the spatial patterns of suppression by PV and SOM neurons were less clustered than a theoretical ordered distance model, and statistically similar to randomly targeted networks, although PV networks tended to be more spatially coherent than SOM networks (Supplementary Fig. 21b–d). Thus, individual PV and SOM neurons do not seem to functionally affect neurons along a distance gradient within their local network.

We then asked whether there was any relationship between the preferred orientation of an inhibitory cell and the preferred orientations of its targets. Comparing the tuning curves of activated inhibitory neurons and the tuning of significantly suppressed cells (Fig. 6b, c), showed that PV cells targeted higher percentages of neurons that matched their own preferred orientations than the orthogonal or the expected percentage predicted by random targeting (Fig. 6d, top panel; preferred orientation (PO), 44.3 ± 7.6%; orthogonal, non-preferred orientation (nonPO), 20.4 ± 4.8%; $P < 0.05$, treating each network as a single observation; $n = 210$ target cells from 4 PV networks; PO versus random PO, $P < 0.05$). However, for SOM cells and networks, the orientation distribution of suppressed cells was more uniform, with no significant difference between the percentage of targeted cells at the preferred orientation versus the orthogonal of the source SOM cells (Fig. 6d, bottom panel; PO, 13.5 ± 3.8%; nonPO, 8.1 ± 4.6%; $P = 0.39$, $n = 238$ target cells from 5 SOM networks). These results indicate that PV cells preferentially target other neurons that have similar preferred
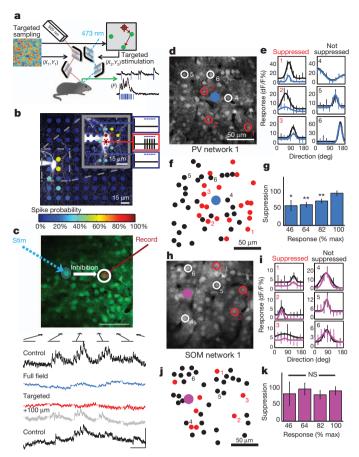
Figure 5 | Dual-laser optical mapping of network connections to reveal maps of functional inhibition by single PV and SOM neurons. **a**, Two laser systems are controlled independently and synchronized for concurrent high-speed imaging and targeted spatial stimulation of a ChR2-expressing inhibitory neuron. **b**, A SOM ChR2+ neuron *in vivo* was patched and recorded while systematically mapping target locations with a focused 473-nm laser; spikes were elicited only when stimulation was targeted on or very near to the soma. The colour bar indicates the spike probability at each location (0–100%). **c**, All-optical circuit mapping. Optical activation of one neuron (blue spot) during targeted recording of another neuron (red spot). Activity in the recorded neuron was measured with no stimulation (Control, black), full-field activation of PV neurons (blue), targeted activation of the PV neuron (red), and while aiming the stimulation beam 100 μm off-target (grey). Visual responses were suppressed by both full-field and single PV-cell activation. **d**, Mapping the influence of targeted PV activation on neighbouring cells revealed some cells that were significantly suppressed by PV activation (red, 1–3) which were clearly intermixed with other nearby cells that were not affected (white, 4–6). **e**, Tuning curves for cells 1–6 comparing cell responses during control (black) or targeted PV-cell activation (blue). **f**, Spatial distribution of all cells in the network that were either significantly suppressed (red) or not (black). **g**, Population data for the amount of PV- and SOM-mediated suppression observed at cells' different response levels across all single cell networks. **h–k**, Same as **d–g**, in experiments in which focal stimulation was targeted to single SOM neurons. *$P < 0.05$; **$P < 0.01$; NS, not significant. Data are shown as mean ± s.e.m.

orientations, whereas SOM neurons seem to have a broader range of targets.

Discussion

Although a growing literature has started to examine the input and firing properties of specific inhibitory neuron classes[19,20,30,31], little is understood about the functional nature of their output. By triggering inhibition and measuring its effects on connected cells in the functioning cortex, we have shown the computational impact of different forms of inhibition on sensory processing. PV neurons principally implement divisive normalization, whereas SOM neurons perform
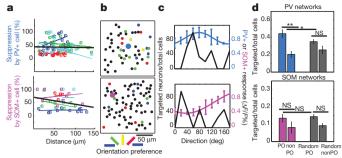
Figure 6 | Spatial and functional analysis of targeting by single PV and SOM neurons. **a**, The amount of suppression in significantly suppressed cells is plotted against their distance from the stimulated PV cell (top) or SOM cell (bottom). Best fit coloured lines show individual networks; black line depicts pooled data. **b**, The orientation preferences of significantly suppressed cells are colour-coded for two example networks (top, PV (blue); bottom, SOM (pink)). **c**, Distribution of preferred orientations of target cells (black lines) for these example networks, superimposed with the orientation tuning curve of the stimulated PV (blue line, top) or SOM (pink line, bottom) neuron. **d**, Bar graphs show the mean proportion of neurons that matched the preferred orientation of stimulated PV neurons (blue bars) or SOM neurons (pink bars), and the orthogonal, non-preferred orientation. Grey bars show the same comparison when the preferred orientations were randomly re-sampled among the neurons in the fields of view (Random PO and Random nonPO). For PV but not SOM, the actual percentage of targets at the PO was greater than expected with random targeting, indicating that PV cells may preferentially suppress cells with similar functional response properties. *$P < 0.05$; **$P < 0.01$; NS, not significant. Data are shown as mean ± s.e.m.

relatively uniform subtraction of responses in their targets, leading to complementary effects on neuronal responses: SOM neurons alter stimulus selectivity, whereas PV neurons preserve selectivity, and PV neurons modulate response gain, whereas SOM neurons shift response levels, leaving response gain unaffected. These effects are mediated by complementary local circuits: PV neurons preferentially target iso-oriented neurons, whereas SOM neurons target cells with a wide range of orientation preferences.

The role of inhibition in shaping stimulus selectivity of visual cortex neurons has been difficult to resolve with previous methods. Intracellular recordings from V1 neurons have found closely matched tuning of excitatory and inhibitory synaptic conductances[8,28,29,32], indicating that inhibition might not sharpen orientation selectivity, and intracellular blockade of inhibition indeed does not seem to affect orientation tuning[33]. However, broadly tuned or untuned inhibition can in principle sharpen neuronal responses; such inhibition has also been described in V1 neurons[10,34,35], and pharmacological blockade of network inhibition broadens orientation selectivity[5–7]. We show that SOM-mediated inhibition targets cells with a range of preferred orientations, and its presence on dendrites probably serves to sharpen the wide orientation preference of dendritic excitatory inputs[36]. Consistent with SOM neurons having this role, genetic reduction of a subset of dendrite-targeting interneurons broadens orientation selectivity[37].

Inhibition has long been proposed to regulate the gain of cortical responses, and we now show that PV cells are crucial for this function. Consistent with our findings, a recent study[26] has demonstrated that activating PV-interneuron populations in mouse visual cortex has a divisive scaling effect on responses of target neurons. Rapid PV-mediated inhibition matched to excitation in time could shape response gain as well as selectivity, as shown in the auditory[9,38], somatosensory[39,40] and prefrontal cortex[41]. Finally, the regulation of response gain by PV cells makes them an attractive mechanism for the developmental regulation of inputs during experience-dependent plasticity of cortical circuits[42].

The distinctive effects of PV and SOM inhibition may arise from their cellular regions of impact[43], and possibly synaptic differences

between the cell types[44,45]. The methods we have described provide a basis for 'functional connectomics' in active cortical circuits, and also reveal the complementary computational roles of specific inhibitory cell classes *in vivo* during sensory processing. In bridging the gap between cellular and network function, these methods should be applicable to many cortical areas and cell types, to elucidate their functional connectivity and embodied computational principles.

METHODS SUMMARY

Details of mice and viral constructs used, animal surgical preparations, *in vitro* slice characterization of ChR2 function, interneuron expression, *in vivo* two-photon guided cell-attached recording, development of high-speed targeted scanning of calcium responses, development of simultaneous two-photon imaging and optogenetic stimulation, the focal stimulation system, and details of data analysis are described in the Supplementary Methods.

1. Freund, T. F. & Buzsáki, G. Interneurons of the hippocampus. *Hippocampus* **6**, 347–470 (1996).
2. Markram, H. *et al.* Interneurons of the neocortical inhibitory system. *Nature Rev. Neurosci.* **5**, 793–807 (2004).
3. Klausberger, T. & Somogyi, P. Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* **321**, 53–57 (2008).
4. Lewis, D. A. GABAergic local circuit neurons and prefrontal cortical dysfunction in schizophrenia. *Brain Res. Brain Res. Rev.* **31**, 270–276 (2000).
5. Sillito, A. M. The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat. *J. Physiol. (Lond.)* **250**, 305–329 (1975).
6. Crook, J. M., Kisvárday, Z. F. & Eysel, U. T. Evidence for a contribution of lateral inhibition to orientation tuning and direction selectivity in cat visual cortex: reversible inactivation of functionally characterized sites combined with neuroanatomical tracing techniques. *Eur. J. Neurosci.* **10**, 2056–2075 (1998).
7. Tsumoto, T., Eckart, W. & Creutzfeldt, O. D. Modification of orientation sensitivity of cat visual cortex neurons by removal of GABA-mediated inhibition. *Exp. Brain Res.* **34**, 351–363 (1979).
8. Ferster, D. & Miller, K. D. Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* **23**, 441–471 (2000).
9. Wehr, M. & Zador, A. M. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* **426**, 442–446 (2003).
10. Monier, C., Chavane, F., Baudot, P., Graham, L. J. & Frégnac, Y. Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. *Neuron* **37**, 663–680 (2003).
11. Carandini, M. & Heeger, D. J. Summation and division by neurons in primate visual cortex. *Science* **264**, 1333–1336 (1994).
12. Somers, D. C., Nelson, S. B. & Sur, M. An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* **15**, 5448–5465 (1995).
13. Reynolds, J. H. & Heeger, D. J. The normalization model of attention. *Neuron* **61**, 168–185 (2009).
14. Ohshiro, T., Angelaki, D. E. & DeAngelis, G. C. A normalization model of multisensory integration. *Nature Neurosci.* **14**, 775–782 (2011).
15. Louie, K. & Glimcher, P. W. Separating value from choice: delay discounting activity in the lateral intraparietal area. *J. Neurosci.* **30**, 5498–5507 (2010).
16. Rudy, B., Fishell, G., Lee, S. & Hjerling-Leffler, J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev. Neurobiol.* **71**, 45–61 (2011).
17. Xu, X., Roby, K. D. & Callaway, E. M. Immunochemical characterization of inhibitory mouse cortical neurons: three chemically distinct classes of inhibitory cells. *J. Comp. Neurol.* **518**, 389–404 (2010).
18. Schummers, J., Yu, H. & Sur, M. Tuned responses of astrocytes and their influence on hemodynamic signals in the visual cortex. *Science* **320**, 1638–1643 (2008).
19. Runyan, C. A. *et al.* Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron* **67**, 847–857 (2010).
20. Bock, D. D. *et al.* Network anatomy and *in vivo* physiology of visual cortical neurons. *Nature* **471**, 177–182 (2011).
21. Yoshimura, Y. & Callaway, E. M. Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nature Neurosci.* **8**, 1552–1559 (2005).
22. Fino, E. & Yuste, R. Dense inhibitory connectivity in neocortex. *Neuron* **69**, 1188–1203 (2011).
23. Packer, A. M. & Yuste, R. Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J. Neurosci.* **31**, 13260–13271 (2011).
24. Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
25. Lillis, K. P., Eng, A., White, J. A. & Mertz, J. Two-photon imaging of spatially extended neuronal network dynamics with high temporal resolution. *J. Neurosci. Methods* **172**, 178–184 (2008).
26. Atallah, B. V., Bruns, W., Carandini, M. & Scanziani, M. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73**, 159–170 (2012).
27. Katzner, S., Busse, L. & Carandini, M. GABAA inhibition controls response gain in visual cortex. *J. Neurosci.* **31**, 5931–5941 (2011).
28. Anderson, J. S., Carandini, M. & Ferster, D. Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *J. Neurophysiol.* **84**, 909–926 (2000).
29. Mariño, J. *et al.* Invariant computations in local cortical networks with balanced excitation and inhibition. *Nature Neurosci.* **8**, 194–201 (2005).
30. Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* **67**, 858–871 (2010).
31. Hofer, S. B. *et al.* Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neurosci.* **14**, 1045–1052 (2011).
32. Tan, A. Y., Brown, B. D., Scholl, B., Mohanty, D. & Priebe, N. J. Orientation selectivity of synaptic input to neurons in mouse and cat primary visual cortex. *J. Neurosci.* **31**, 12339–12350 (2011).
33. Nelson, S., Toth, L., Sheth, B. & Sur, M. Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science* **265**, 774–777 (1994).
34. Xing, D., Ringach, D. L., Hawken, M. J. & Shapley, R. M. Untuned suppression makes a major contribution to the enhancement of orientation selectivity in macaque v1. *J. Neurosci.* **31**, 15972–15982 (2011).
35. Liu, B. H. *et al.* Broad inhibition sharpens orientation selectivity by expanding input dynamic range in mouse simple cells. *Neuron* **71**, 542–554 (2011).
36. Jia, H., Rochefort, N. L., Chen, X. & Konnerth, A. Dendritic organization of sensory input to cortical neurons *in vivo*. *Nature* **464**, 1307–1312 (2010).
37. Mao, R. *et al.* Influence of a subtype of inhibitory interneuron on stimulus-specific responses in visual cortex. *Cereb. Cortex* **22**, 493–508 (2011).
38. Tan, A. Y., Zhang, L. I., Merzenich, M. M. & Schreiner, C. E. Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. *J. Neurophysiol.* **92**, 630–643 (2004).
39. Okun, M. & Lampl, I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neurosci.* **11**, 535–537 (2008).
40. Gentet, L. J., Avermann, M., Matyas, F., Staiger, J. F. & Petersen, C. C. Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* **65**, 422–435 (2010).
41. Haider, B., Duque, A., Hasenstaub, A. R. & McCormick, D. A. Neocortical network activity *in vivo* is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* **26**, 4535–4545 (2006).
42. Yazaki-Sugiyama, Y., Kang, S., Câteau, H., Fukai, T. & Hensch, T. K. Bidirectional plasticity in fast-spiking GABA circuits by visual experience. *Nature* **462**, 218–221 (2009).
43. Koch, C., Poggio, T. & Torre, V. Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proc. Natl Acad. Sci. USA* **80**, 2799–2802 (1983).
44. Kulik, A. *et al.* Compartment-dependent colocalization of Kir3.2-containing K$^+$ channels and GABAB receptors in hippocampal pyramidal cells. *J. Neurosci.* **26**, 4289–4297 (2006).
45. Nusser, Z., Sieghart, W., Benke, D., Fritschy, J. M. & Somogyi, P. Differential synaptic localization of two major gamma-aminobutyric acid type A receptor alpha subunits on hippocampal pyramidal cells. *Proc. Natl Acad. Sci. USA* **93**, 11939–11944 (1996).

Author Contributions N.R.W. conceived experiments, designed and engineered circuit interface and analysis systems, carried out *in vivo* and *in vitro* experiments, and performed analyses. C.A.R. conceived experiments, performed surgeries and viral injections, carried out *in vivo* experiments, and performed analyses. F.L.W. carried out *in vivo* experiments, and performed analyses. M.S. conceived experiments and contributed to analysis of experiments. N.R.W., C.A.R. and M.S. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.S. (msur@mit.edu) or N.R.W. (nathan1@mit.edu).

LETTER

A massive, cooling–flow–induced starburst in the core of a luminous cluster of galaxies

M. McDonald[1], M. Bayliss[2], B. A. Benson[3,4], R. J. Foley[5], J. Ruel[2], P. Sullivan[1], S. Veilleux[6,7], K. A. Aird[8], M. L. N. Ashby[5], M. Bautz[1], G. Bazin[9,10], L. E. Bleem[3,11], M. Brodwin[12], J. E. Carlstrom[3,4,11,13,14], C. L. Chang[3,4,14], H. M. Cho[15], A. Clocchiatti[16], T. M. Crawford[3,13], A. T. Crites[3,13], T. de Haan[17], S. Desai[9,10], M. A. Dobbs[17], J. P. Dudley[17], E. Egami[18], W. R. Forman[5], G. P. Garmire[19], E. M. George[20], M. D. Gladders[3,13], A. H. Gonzalez[21], N. W. Halverson[22], N. L. Harrington[20], F. W. High[3,13], G. P. Holder[17], W. L. Holzapfel[20], S. Hoover[3,4], J. D. Hrubes[8], C. Jones[5], M. Joy[23], R. Keisler[3,11], L. Knox[24], A. T. Lee[19,25], E. M. Leitch[3,13], J. Liu[9,10], M. Lueker[19,26], D. Luong-Van[8], A. Mantz[3], D. P. Marrone[18], J. J. McMahon[3,4,27], J. Mehl[3,13], S. S. Meyer[3,4,11,13], E. D. Miller[1], L. Mocanu[3,13], J. J. Mohr[9,10,28], T. E. Montroy[29], S. S. Murray[5], T. Natoli[3,11], S. Padin[3,13,26], T. Plagge[3,13], C. Pryke[30], T. D. Rawle[18], C. L. Reichardt[20], A. Rest[31], M. Rex[18], J. E. Ruhl[29], B. R. Saliwanchik[29], A. Saro[9,10], J. T. Sayre[29], K. K. Schaffer[3,4,32], L. Shaw[17,33], E. Shirokoff[20,25], R. Simcoe[1], J. Song[27], H. G. Spieler[25], B. Stalder[5], Z. Staniszewski[29], A. A. Stark[5], K. Story[3,11], C. W. Stubbs[2,5], R. Šuhada[9], A. van Engelen[17], K. Vanderlinde[17], J. D. Vieira[3,11,26], A. Vikhlinin[5], R. Williamson[3,13], O. Zahn[20,34,35] & A. Zenteno[9,10]

In the cores of some clusters of galaxies the hot intracluster plasma is dense enough that it should cool radiatively in the cluster's lifetime[1–3], leading to continuous 'cooling flows' of gas sinking towards the cluster centre, yet no such cooling flow has been observed. The low observed star-formation rates[4,5] and cool gas masses[6] for these 'cool-core' clusters suggest that much of the cooling must be offset by feedback to prevent the formation of a runaway cooling flow[7–10]. Here we report X-ray, optical and infrared observations of the galaxy cluster SPT-CLJ2344-4243 (ref. 11) at redshift $z = 0.596$. These observations reveal an exceptionally luminous ($8.2 \times 10^{45}\,\mathrm{erg\,s^{-1}}$) galaxy cluster that hosts an extremely strong cooling flow (around 3,820 solar masses a year). Further, the central galaxy in this cluster appears to be experiencing a massive starburst (formation of around 740 solar masses a year), which suggests that the feedback source responsible for preventing runaway cooling in nearby cool-core clusters may not yet be fully established in SPT-CLJ2344-4243. This large star-formation rate implies that a significant fraction of the stars in the central galaxy of this cluster may form through accretion of the intracluster medium, rather than (as is currently thought) assembling entirely via mergers.

The galaxy cluster SPT-CLJ2344-4243 was discovered by the South Pole Telescope[12] via the Sunyaev–Zel'dovich effect, with an initial estimated[11] mass of $16.6 \times 10^{14}\,\mathrm{M_\odot}$, where $\mathrm{M_\odot}$ is the mass of the Sun. These data were supplemented with new broadband optical g, r, i, z imaging from the Mosaic II camera on the Blanco 4-m telescope (Fig. 1), optical spectroscopy from the Gemini Multi-Object Spectrograph

(GMOS) on the Gemini South 8.1-m telescope, optical long-slit spectroscopy using the Inamori-Magellan Areal Camera and Spectrograph (IMACS) on the 6.5-m Magellan telescope, near-infrared long-slit spectroscopy using the Folded-port InfraRed Echellette (FIRE) on the 6.5-m Magellan telescope, mid–far-infrared imaging using the Photodetector Array Camera and Spectrometer (PACS) and the Spectral and Photometric Imaging Receiver (SPIRE) on the Herschel Space Observatory, and X-ray imaging spectroscopy using the ACIS-I camera on the Chandra X-ray Observatory. Additionally, we have acquired archival near–far-ultraviolet imaging from the orbiting space telescope Galaxy Evolution Explorer (GALEX) archives, near–mid-infrared imaging from the Two Micron All-Sky Survey (2MASS) and Wide-field Infrared Survey Explorer (WISE) archives, and 843-MHz radio imaging from the Sydney University Molonglo Sky Survey (SUMSS) survey. Further details of these data and their processing can be found in the Supplementary Information.

We estimate the mass of SPT-CLJ2344-4243 from the X-ray-measured pressure ($Y_X \equiv M_{gas} \times T_X$, where T_X is the X-ray-measured temperature of the intracluster gas and M_{gas} is its mass) of the intracluster medium, using an externally calibrated pressure–mass (Y_X–M) relation. The relation was calibrated using a local sample of relaxed (that is, non-merging) clusters from X-ray estimates of the total mass that assumed hydrostatic equilibrium[13]. By iteratively adjusting the value of r_{500} (where r_{500} (or r_{200}) is the radius for which the enclosed average density is 500 (or 200) times the critical (or average) density of the Universe) such that the Y_X–M_{500} relation is satisfied, we converge on values of $r_{500} = 1.3\,\mathrm{Mpc}$
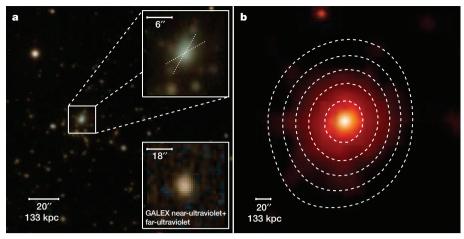
Figure 1 | False-colour images of the galaxies and intracluster plasma in the galaxy cluster SPT-CLJ2344-4243. a, This colour-composite image of SPT-CLJ2344-4243 is based on a red–green–blue combination of optical *r*, *i*, *z* images. Galaxies that make up the galaxy cluster share a common brown colour, owing to their similar star-formation histories and common distance from Earth. The central galaxy, which is both the most massive and most luminous galaxy in the cluster, is considerably bluer than the rest of the member galaxies, suggesting significantly younger stellar populations. This is more obvious in the zoomed-in inset. The lower right inset, which shows an ultraviolet colour-composite, reveals a bright ultraviolet source, with no accompanying emission from the surrounding member galaxies. Dotted lines represent the orientation of the optical and near-infrared long-slit spectra. **b**, This false-colour image shows the adaptively smoothed X-ray data, with photon energies from 0.7–2.0 keV (to minimize AGN contribution), of SPT-CLJ2344-4243. The image clearly shows the luminous, centrally concentrated core, as well as the relatively smooth, relaxed morphology of the intracluster medium. White contours represent the Sunyaev–Zel'dovich decrement (significance levels of 5, 10, 15, 20 and 25) against the cosmic microwave background. The circularity of these contours agree with the scenario that this system is not currently undergoing a major merger with another galaxy cluster.

and $M_{500, Y_x} = 12.6^{+2.0}_{-1.5} \times 10^{14} M_\odot$. At r_{200}, this corresponds to $M_{200,Y_x} \approx 25 \times 10^{14} M_\odot$, which makes SPT-CLJ2344-4243 amongst the most massive known clusters in the Universe[14,15]. The GMOS multi-object spectroscopy of 26 galaxies, each exhibiting only absorption features, yielded a robust biweight estimate of the redshift ($z = 0.596 \pm 0.002$) and velocity dispersion ($1,700^{+300}_{-200}$ km s^{-1}), the latter being consistent with the picture of an extremely massive cluster. The velocity distribution is consistent with a Gaussian distribution, but the limited number of redshifts does not preclude velocity substructure or multimodality. The smooth X-ray isophotes suggest that the cluster may be relaxed, and although the cluster member velocity distribution is consistent with an undisturbed cluster, the velocity data lack the statistical power to constrain the cluster's dynamical state robustly.

The integrated rest-frame X-ray luminosity at 2–10 keV, $L_{2-10 \, keV} = 8.2^{+0.1}_{-0.2} \times 10^{45}$ erg s^{-1} within r_{500}, is greater than any other known cluster in this band. The high central luminosity, which is predominantly cooling radiation, in turn results in a high X-ray cooling rate, as defined by $\frac{dM}{dt} = \frac{2L\mu m_p}{5kT}$, where μ is the mean molecular mass of the intracluster medium, L is the X-ray luminosity, m_p is the mass of a proton and kT is the temperature of the intracluster gas. Assuming a cooling radius of 100 kpc (see Supplementary Information), we measure an intracluster medium cooling rate of $3,820 \pm 530 M_\odot$ yr^{-1}, making this the strongest cooling flow yet discovered (see Table 1 for comparison to other clusters). The intracluster medium in SPT-CLJ2344-4243 exhibits a significant drop in temperature, accompanied by a rise in the metallicity in the central 100 kpc that is reminiscent of nearby cool core clusters. Furthermore, the short

central cooling time (less than a billion years, <1 Gyr), along with the low central entropy (<100 keV cm^2), resembles nearby strongly cooling cores, such as the Perseus[16] and PKS0745-191 (ref. 17) clusters. The discovery of a strong cool core at $z = 0.596$ is particularly remarkable because recent X-ray and optical surveys have found a general lack of strong cool cores at $z > 0.4$ (refs 18–20), with relatively few exceptions.

Much like the central galaxies of low-z cool-core clusters[21–23], SPT-CLJ2344-4243 exhibits bright, spatially extended, optical line emission (that is, [O II], Hβ, [O III], [O I], Hα, [N II], [S II], and so on; Fig. 2). We were fortunate to intersect what appears to be an extended filament with one of our randomly oriented slits, which has a length of about 70 kpc. This is similar in extent to the most extended optical filaments in the core of the Perseus cluster[24], and orders of magnitude larger than typical jets in cluster cores (such as M87; ref. 25). The diagnostic line ratios [N II]/Hα, [S II]/Hα, [O I]/Hα, [O III]/Hβ, and [O III]/[O II] show evidence for a Seyfert-like active galactic nucleus (AGN) in the central galaxy, while at radii exceeding 2″, the optical line ratios resemble those in the star-forming filaments of $z \approx 0$ cool-core clusters[26] (see Supplementary Information).

Apart from the exceptionally high X-ray luminosity and central cooling rate, what sets this system apart from the majority of nearby galaxy clusters is that there is significant evidence for a dusty starburst in the central galaxy of SPT-CLJ2344-4243. The rest-frame 0.1–500 μm spectral energy distribution (Fig. 3) of the central galaxy most closely resembles that of an ultraluminous, infrared galaxy; such galaxies are known to have heavily obscured starbursts (about $200-1,000 M_\odot$ yr^{-1}) and central AGNs. This scenario is corroborated by our observation of significant Balmer reddening ($E(B - V)_{global} \approx 0.3$)

Table 1 | Well-studied, strong, cool-core clusters compared to SPT-CLJ2344-4243

Cluster	z	$L_{2.0-10.0 \, keV}$ (10^{44} erg s^{-1})	kT (keV)	dM/dt (M$_\odot$ yr^{-1})	SFR (M$_\odot$ yr^{-1})	ε_{cool}
Perseus	0.0179	11	5.5	556	37	0.07
PKS0745-191	0.1028	29.5	6.71	1,455	20	0.01
Zw 3146	0.2906	36.9	6.4	2,228	79	0.04
RX J1347.5-1145	0.451	60	10.0	1,900	23	0.01
SPT-CLJ2344-4243	0.596	82^{+1}_{-2}	$13.0^{+2.4}_{-3.4}$	3,820 ± 530	740 ± 160	0.19 ± 0.05

Prior to this work, RX J1347.5-1145 was considered both the most X-ray luminous and strongest cooling galaxy cluster, with a luminosity of $L_{2-10 \, keV} = 60 \times 10^{44}$ erg s^{-1} and cooling rate of 1,900M$_\odot$ yr^{-1}. The star-formation rate of the central galaxy in SPT-CLJ2344-4243 is exceptionally high. We quantify the efficiency of converting the cooling flow into stars with the parameter ε_{cool}, which is simply the star-formation rate normalized to the classical cooling rate. The high star-formation rate implies that SPT-CLJ2344-4243 is converting around 20% of the cooling flow into stars, which is considerably higher than the vast majority of low-redshift cool-core clusters. X-ray properties and star-formation rates of the lower-redshift clusters are taken from the literature[6,28,29].
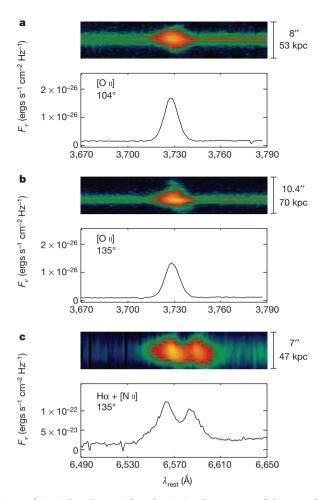
Figure 2 | Optical and near-infrared emission-line spectra of the central galaxy in SPT-CLJ2344-4243. The false-colour images highlight the extended line emission ([O II] in **a** and **b**, Hα+[N II] in **c**), where the vertical axis is the spatial direction (along the slit) and the horizontal axis is the spectral direction. F_ν is the emission flux. These emission lines result from warm, ionized gas at about 10^4 K, which is probably heated by a combination of young stars, shocks and feedback from the central active black hole. The fact that the [O II] emission is significantly more extended in one direction (position angle of the slit, relative to north $\theta \approx 135°$) suggests that the emission is non-axisymmetric, and is consistent with the scenario of radial line-emitting filaments. The extent of this emission (>50 kpc) is reminiscent of optical filaments observed in the core of the Perseus cluster[24]. Beneath each colour image, we show the spectrum that is generated by summing along columns of the colour image. These spectra show the high signal-to-noise ratio of these emission lines, leading to high-confidence estimates of the emission line luminosity. **c**, The near-infrared spectrum (in the observed frame, $\lambda_{H\alpha} = 1.05\,\mu m$) shows emission from both the Hα and [N II] lines, extended over similar radii (>50 kpc).

and strong 24–160 μm emission, combined with signatures of ongoing star formation (near- and far-ultraviolet emission, bright nebular emission lines, weak 4,000-Å break) and a heavily obscured central AGN ($E(B - V)_{nuclear} \approx 0.5$, $n_{H, X-ray} \approx 40 \times 10^{22}\,cm^{-2}$). Using the full multi-wavelength data set, which includes X-ray, near–far ultraviolet, optical, near–far infrared, and radio data, we estimate an extinction-corrected, AGN-subtracted star-formation rate of $740 \pm 160 M_\odot\,yr^{-1}$, assuming a geometric correction of 45% for the long-slit spectroscopy and an AGN contamination fraction of around 40%–50% (see Supplementary Information for details).

The presence of extended (around 70 kpc), morphologically complex (Fig. 2), star-forming filaments coincident with the central galaxy in SPT-CLJ2344-4243 is reminiscent of low-z cool-core clusters like Perseus and PKS0745-191. However, while these clusters have substantial amounts of star formation (around 20 solar masses a
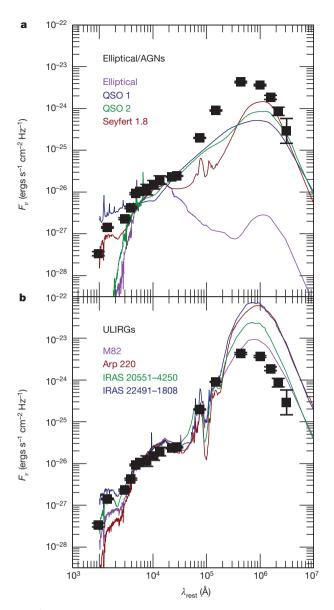
Figure 3 | Spectral energy distribution of the central galaxy in SPT-CLJ2344-4243 from the far-ultraviolet to the far-infrared. We show on the y-axis the integrated specific flux (F_ν) as a function of wavelength in the rest frame of the galaxy cluster, along with the associated 1σ measurement error. A typical central cluster galaxy is morphologically classified as an 'elliptical' galaxy, and has an average spectral energy distribution like that shown in purple in **a**. In contrast, the central galaxy in SPT-CLJ2344-4243 has a considerable excess of emission at both ultraviolet and infrared wavelengths, indicative of strong star formation. While there is significant evidence for strong feedback from the central, supermassive black hole in this cluster (see Supplementary Information), this figure shows that simple models of AGNs (quasars type 1 and 2, and Seyfert type 1.8 shown in **a**) are unable to reproduce the spectral shape of the central galaxy in SPT-CLJ2344-4243. However, in **b** we show templates[30] of four different dusty starbursts, or ultra-luminous infrared galaxies (ULIRGs), which provide a much better match to the data. This good agreement suggests that the central galaxy in SPT-CLJ2344-4243, unlike typical central cluster galaxies, contains a dusty starburst and a heavily obscured AGN. Specifically, the spectral shape is most similar to those of M82, a dusty starburst with a strong wind, and IRAS 20551-4250, which is a composite of a highly obscured AGN and a vigorous, dusty starburst.

year)[4,5], this is still orders of magnitude less than predicted by the classical cooling estimates based on the X-ray luminosity (Table 1). This disagreement has become known as the 'cooling flow' problem, and it is generally assumed that some form of feedback is responsible for halting the cooling intracluster medium before it reaches the cold

phase. SPT-CLJ2344-4243, however, represents one of a very few exceptions to this general trend[4,27], where the high star-formation rate represents a significant fraction of the massive cooling flow (Table 1). Whatever feedback mechanism is responsible for preventing runaway cooling of the intracluster medium in low-redshift galaxy clusters is clearly operating with a lower efficiency in this system. Although the central galaxy hosts an AGN, as shown by a hard X-ray point source and strong radio emission (see Supplementary Information), it may be that we are observing this system during a small window in time when the AGN is rapidly feeding off the cooling flow, but the power output of the AGN has not yet fully coupled to the intracluster medium, and therefore is able to halt a smaller fraction of the total cooling than in typical low-redshift clusters (such as Perseus). That systems with such high cooling and star-formation rates are not observed at $z = 0$ suggests that either this system is unique, or that the mechanism which quenches cooling may have been less effective in the early Universe. Further studies of distant, strongly cooling galaxy clusters are needed to differentiate between these two scenarios.

The high star-formation rates inferred from optical line emission and near–far-ultraviolet, optical, and mid–far-infrared continuum emission, combined with the strong signatures of X-ray cooling, suggest that the central galaxy in SPT-CLJ2344-4243 may form a substantial fraction of its stars through an intense, short-lived cooling phase of the intracluster medium. Such strong cooling cannot be sustained for a significant amount of time, or both the central galaxy and its supermassive black hole would become too massive, and the central galaxy would have stellar populations considerably younger than those observed in giant elliptical galaxies today. This implies that episodes of strong cooling are short-lived, in contrast to the longer episodes of strong feedback observed in nearby clusters.

1. Lea, S. M., Silk, J., Kellogg, E. & Murray, S. Thermal-Bremsstrahlung interpretation of cluster X-ray sources. *Astrophys. J.* **184**, L105 (1973).
2. Cowie, L. L. & Binney, J. Radiative regulation of gas flow within clusters of galaxies—a model for cluster X-ray sources. *Astrophys. J.* **215**, 723–732 (1977).
3. Fabian, A. C. & Nulsen, P. E. J. Subsonic accretion of cooling gas in clusters of galaxies. *Mon. Not. R. Astron. Soc.* **180**, 479–484 (1977).
4. O'Dea, C. P. *et al.* An infrared survey of brightest cluster galaxies. II. Why are some brightest cluster galaxies forming stars? *Astrophys. J.* **681**, 1035–1045 (2008).
5. McDonald, M., Veilleux, S., Rupke, D. S. N., Mushotzky, R. & Reynolds, C. Star formation efficiency in the cool cores of galaxy clusters. *Astrophys. J.* **734**, 95 (2011).
6. Edge, A. C. The detection of molecular gas in the central galaxies of cooling flow clusters. *Mon. Not. R. Astron. Soc.* **328**, 762–782 (2001).
7. McNamara, B. R. & Nulsen, P. E. J. Heating hot atmospheres with active galactic nuclei. *Annu. Rev. Astron. Astrophys.* **45**, 117–175 (2007).
8. Fabian, A. C. Observational evidence of AGN feedback. *Annu. Rev. Astron. Astrophys.* (in the press); preprint at http://arxiv.org/abs/1204.4114 (2012).
9. Mathews, W. G. Stopping cooling flows with cosmic-ray feedback. *Astrophys. J. Lett.* **695**, L49–L52 (2009).
10. Gómez, P. L., Loken, C., Roettiger, K. & Burns, J. O. Do cooling flows survive cluster mergers? *Astrophys. J.* **569**, 122–133 (2002).
11. Williamson, R. *et al.* A Sunyaev-Zel'dovich-selected sample of the most massive galaxy clusters in the 2500 deg^2 South Pole Telescope Survey. *Astrophys. J.* **738**, 139 (2011).
12. Carlstrom, J. E. *et al.* The 10 meter South Pole Telescope. *Proc. Astron. Soc. Pacif.* **123**, 568–581 (2011).
13. Vikhlinin, A. *et al.* Chandra Cluster Cosmology Project. II. Samples and X-ray data reduction. *Astrophys. J.* **692**, 1033–1059 (2009).
14. Menanteau, F. *et al.* The Atacama Cosmology Telescope: ACT-CL J0102–4915 "El Gordo", a massive merging cluster at redshift 0.87. *Astrophys. J.* **748**, 7 (2011).
15. Foley, R. J. *et al.* Discovery and cosmological implications of SPT-CL J2106–5844, the most massive known cluster at z>1. *Astrophys. J.* **731**, 86 (2011).
16. Fabian, A. C. *et al.* Chandra imaging of the complex X-ray core of the Perseus cluster. *Mon. Not. R. Astron. Soc.* **318**, L65–L68 (2000).
17. Allen, S. W., Fabian, A. C. & Kneib, J. P. A combined X-ray and gravitational lensing study of the massive cooling-flow cluster PKS 0745–191. *Mon. Not. R. Astron. Soc.* **279**, 615–635 (1996).
18. Vikhlinin, A. *et al.* in *Heating versus Cooling in Galaxies and Clusters of Galaxies* (eds Böhringer, H., Pratt, G. W., Finoguenov, A. & Schuecker, P.) 48 (2007); preprint at http://arxiv.org/astro-ph/0611438..
19. Santos, J. S. *et al.* Searching for cool core clusters at high redshift. *Astron. Astrophys.* **483**, 35–47 (2008).
20. McDonald, M. Optical line emission in brightest cluster galaxies at 0 < z 0.6: evidence for a lack of strong cool cores 3.5 Gyr ago? *Astrophys. J. Lett.* **742**, L35 (2011).
21. Hu, E. M., Cowie, L. L. & Wang, Z. Long-slit spectroscopy of gas in the cores of X-ray luminous clusters. *Astrophys. J.* **59** (Suppl.), 447–498 (1985).
22. Heckman, T. M., Baum, S. A., van Breugel, W. J. M. & McCarthy, P. Dynamical, physical, and chemical properties of emission-line nebulae in cooling flows. *Astrophys. J.* **338**, 48–77 (1989).
23. McDonald, M., Veilleux, S., Rupke, D. S. N. & Mushotzky, R. On the origin of the extended Hα filaments in cooling flow clusters. *Astrophys. J.* **721**, 1262–1283 (2010).
24. Conselice, C. J., Gallagher, J. S. III & Wyse, R. F. G. On the nature of the NGC 1275 system. *Astron. J.* **122**, 2281–2300 (2001).
25. Sparks, W. B., Biretta, J. A. & Macchetto, F. The jet of M87 at tenth-arcsecond resolution: optical, ultraviolet, and radio observations. *Astrophys. J.* **473**, 254 (1996).
26. McDonald, M., Veilleux, S. & Rupke, D. S. N. Optical spectroscopy of Hα filaments in cool core clusters: kinematics, reddening, and sources of ionization. *Astrophys. J.* **746**, 153 (2012).
27. McNamara, B. R. *et al.* The starburst in the Abell 1835 Cluster Central Galaxy: a case study of galaxy formation regulated by an outburst from a supermassive black hole. *Astrophys. J.* **648**, 164–175 (2006).
28. Allen, S. W. The properties of cooling flows in X-ray luminous clusters of galaxies. *Mon. Not. R. Astron. Soc.* **315**, 269–295 (2000).
29. Gitti, M. & Schindler, S. XMM-Newton observation of the most X-ray-luminous galaxy cluster RX J1347.5–1145. *Astron. Astrophys.* **427**, L9–L12 (2004).
30. Polletta, M. *et al.* Spectral energy distributions of hard X-ray selected active galactic nuclei in the XMM-Newton Medium Deep Survey. *Astrophys. J.* **663**, 81–102 (2007).

Author Contributions M.M.D. reduced the X-ray and optical long slit spectroscopic data, performed the main analysis, and wrote the paper, with significant assistance from B.A.B., R.J.F. and S.V., and comments from all other authors. M.B. and J.R. reduced multi-slit observations of SPT-CLJ2344-4243 and performed the velocity dispersion analysis. P.S. and R.S. obtained the infrared spectroscopy, and P.S. reduced these data. All other authors (listed alphabetically) have contributed as part of the South Pole Telescope collaboration, by either their involvement with the initial cluster discovery with the South Pole Telescope and/or multi-wavelength follow-up.

Room–temperature solid–state maser

Mark Oxborrow[1], Jonathan D. Breeze[2] & Neil M. Alford[2]

The invention of the laser has resulted in many innovations, and the device has become ubiquitous. However, the maser, which amplifies microwave radiation rather than visible light, has not had as large an impact, despite being instrumental in the laser's birth[1,2]. The maser's relative obscurity has mainly been due to the inconvenience of the operating conditions needed for its various realizations: atomic[3] and free-electron[4] masers require vacuum chambers and pumping; and solid-state masers[5], although they excel as low-noise amplifiers[6] and are occasionally incorporated in ultrastable oscillators[7,8], typically require cryogenic refrigeration. Most realizations of masers also require strong magnets, magnetic shielding or both. Overcoming these various obstacles would pave the way for improvements such as more-sensitive chemical assays, more-precise determinations of biomolecular structure and function, and more-accurate medical diagnostics (including tomography) based on enhanced magnetic resonance spectrometers[9] incorporating maser amplifiers and oscillators. Here we report the experimental demonstration of a solid-state maser operating at room temperature in pulsed mode. It works on a laboratory bench, in air, in the terrestrial magnetic field and amplifies at around 1.45 gigahertz. In contrast to the cryogenic ruby maser[6], in our maser the gain medium is an organic mixed molecular crystal, _p_-terphenyl doped with pentacene, the latter being photo-excited by yellow light. The maser's pumping mechanism exploits spin-selective molecular intersystem crossing[10] into pentacene's triplet ground state[11,12]. When configured as an oscillator, the solid-state maser's measured output power of around −10 decibel milliwatts is approximately 100 million times greater than that of an atomic hydrogen maser[3], which oscillates at a similar frequency (about 1.42 gigahertz). By exploiting the high levels of spin polarization readily generated by intersystem crossing in photo-excited pentacene and other aromatic molecules, this new type of maser seems to be capable of amplifying with a residual noise temperature far below room temperature.

To the best of our knowledge, no-one has previously attempted to make a zero-field maser[13] out of an optically pumped crystal of pentacene-doped _p_-terphenyl. Two further novel features of the device we report here are a microwave cavity containing a dielectric ring of sapphire, which supports a high-Q electromagnetic mode (where Q is the mode's quality factor), and a medical laser, designed for the treatment of vascular skin lesions, that is here used to deliver intense pulses of pump light into the pentacene:p-terphenyl maser crystal at its optimal absorption wavelength (∼585 nm).

Because they are key components within many measurement and telecommunication devices and systems, there has been intensive research on low-noise amplifiers for the purpose of improving their performance and reducing their operating requirements[14]. As low-noise amplifiers, conventional solid-state masers[5] offer low residual noise temperatures[6] and, in contrast to microfabricated devices, are extremely rugged with respect to electromagnetic shock and overloading. Furthermore, they have low intermodulation distortion[6], so reducing interference from out-of-band signals, and low flicker noise[7], thus enabling precise measurements. The active material within a conventional solid-state maser is a dielectric crystal containing paramagnetic ions, for example ruby, that is maintained typically below 10 K and exposed (with exceptions[8,13]) to a strong d.c. magnetic bias field. Despite the inconvenience of these operating conditions, as well as competition from amplifiers based on cooled high-electron-mobility transistors (HEMTs)[15], ruby masers are used in deep-space communications[6]. They were used, for example, by NASA to receive images of the Solar System's outer planets and their moons from the Voyager space probes[16].

The lowest-noise semiconductor-based amplifiers at microwave frequencies, which today most often incorporate cooled HEMTs[15], suffer from a particular limitation. On cooling, the HEMT amplifier's residual noise temperature, although low enough for many applications, is difficult to reduce below a few kelvin[17]. Although this 'HEMT plateau' can be lowered somewhat by using lower device powers[18], it has motivated the successful development of amplifiers based on resonant superconducting interference devices[17], which offer substantially lower noise temperatures, albeit over more limited bandwidths. But with both HEMTs and such devices, the attainment of millikelvin noise temperatures demands subkelvin refrigeration, as supplied by a dilution refrigerator or its equivalent.

It is known why conventional solid-state masers[5,6] cannot be made to work at room temperature[19]. The most serious problem is that the rate of spin–lattice relaxation, $1/T_1$, and, thus, the microwave power required to saturate the maser's pump transition and, in turn, the required thermal cooling power of the maser's refrigerator, increase extremely rapidly with the absolute (lattice) temperature, T, of the maser crystal. For two-phonon (Raman) scattering off Kramers-type paramagnetic ions in a three-dimensional lattice well below its Debye temperature, this relaxation rate scales as T^9. A second problem is that, even if the maser's pump transition can be saturated by pumping hard, the population inversion and, hence, the maser crystal's gain scale as $I \times hf_{\mathrm{mas.}}/kT$, where $I \equiv f_{\mathrm{pump}}/2f_{\mathrm{mas.}} - 1$ is the inversion ratio[5], $f_{\mathrm{mas.}}$
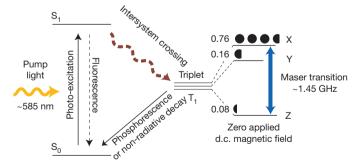

Figure 1 | Maser pumping scheme (Jablonski diagram). Pentacene guest molecules within a _p_-terphenyl host lattice are driven from their singlet ground states, S_0, into their first-excited single states, S_1, by absorbing photons of yellow pump light. From S_1, they predominantly undergo intersystem crossing into their triplet ground states, T_1. This process is spin selective; the relative population rates[11,12] (splitting ratios) into the three spin sublevels of T_1 are as stated (also represented by solid black circles). As can be seen, the uppermost sublevel, X, is preferentially populated, resulting in a strong initial population inversion between it and the lowest sublevel, Z, so providing the conditions for masing through the X ↔ Z transition.

[1]National Physical Laboratory, Hampton Road, Teddington TW11 0LW, UK. [2]Imperial College, Department of Materials, South Kensington Campus, London SW7 2AZ, UK.

and f_{pump} are respectively the maser signal and pump frequencies, and k and h are respectively Boltzmann's and Planck's constants. The residual noise temperature of a conventionally pumped solid-state maser amplifier thus scales as T, so nullifying the maser's low-noise advantage at higher operating temperatures. It was realized very early on that, at the expense of energy efficiency, larger population inversions and, thus, lower noise temperatures and somewhat higher sustainable operating temperatures can be obtained, for a fixed $f_{mas.}$, by increasing f_{pump}—in the extreme case by driving optical pump transitions with light[20], which can often be supplied by a laser. But optical pumping alone does not obviate the first problem, that is, the rapid increase in the rate of spin–lattice relaxation with temperature[21].

Steps have been taken towards the realization of solid-state masers at room temperature using optically excited organic molecules in strong d.c. magnetic fields, focusing largely on C_{60} and porphyrins known to have long spin-polarization lifetimes[19] (T_1). But the 'triplet mechanism', of which Fig. 1 is an example, was considered unsuitable for maser operation at room temperature[22]. Maser threshold was almost reached by instead using the more complex, indirect, radical-triplet-pair mechanism (the triplet state in question being that of etioporphyrin); see fig. 3 of ref. 23. More recently, the potential of various triplet-mechanism photochemistries, including the fullerene derivative PCBM in polystyrene, was assessed in the context of realizing particle acceleration by stimulated emission of radiation at X-band microwave frequencies[24].

The maser reported here (Figs 1 and 2) is based on the triplet mechanism. It exploits spin-selective intersystem crossing[10] into the triplet ground state of pentacene molecules doping crystalline p-terphenyl[11], as has been studied in connection with realizations of both dynamic nuclear polarization[25] and electron spin resonance of single molecules[12]. On photo-excitation from pentacene's ground state, S_0, to its first excited singlet state, S_1, with yellow pump light (Fig. 1), intersystem crossing provides a strongly inverted spin population between the uppermost (X) and lowest (Z) spin sublevels of pentacene's triplet ground state[11,12]. The relatively high zero-field splitting (as compared with PCBM, for example) makes it practicable to use a microwave resonator with no applied d.c. magnetic field, which greatly simplifies the experiment set-up. Our chosen design of resonator (Fig. 2) is cylindrical and supports a frequency-tunable transverse-electric mode, $TE_{01\delta}$, which has no azimuthal dependence,

has a single antinode in the radial and the axial directions, and has a mode volume of $\sim 50\,cm^3$. With the pentacene:p-terphenyl crystal in place, this mode's unloaded quality factor (Q) was measured to be around 2×10^5. When critically coupled through a magnetic loop probe in the side of the resonator's enclosing metal can (Fig. 2), the mode offered a field-to-power conversion factor of $\Lambda \approx 5\,G\,W^{-1/2}$, which is proportional to the square root of the mode's magnetic Purcell factor[26].

An unpolarized, ~ 10-mm-diameter beam of yellow (585 ± 5 nm) pump light from a Candela SPTL-1b long-pulsed rhodamine-6G dye laser was directed at the maser crystal. This laser was capable of delivering up to 4 J of energy in a single pulse, at a rate of one pulse per second, although pulses of only 0.5 J were used in our experiment. (This pulse energy, the time-integrated power of the yellow trace in Fig. 3, was confirmed with a Molectron JMAX 43 pyroelectric joulemeter.) We estimate that ~ 0.3 J was absorbed in the maser crystal per pulse. With the frequency of the $TE_{01\delta}$ mode set near to the centre of the X$\leftrightarrow$Z spin transition, each pump pulse absorbed by the crystal produced a vigorous burst of maser oscillation (Fig. 3). The oscillation's magnitude as a function of the $TE_{01\delta}$ mode's frequency is shown in Fig. 4. The instantaneous output power of the maser burst (Fig. 3, blue trace), as monitored directly by a microwave receiver in the form of an HP8593E spectrum analyser set to fixed-frequency mode, was observed to exceed -10 dBm. Over 1,000 pump pulses, no change in the size or shape of the maser burst was noticeable provided that the resonator was kept in tune. By inserting neutral-density filters into the pump beam to extinguish the maser oscillation, the threshold pump energy exiting from the laser was determined to be 80 mJ over an effective pump pulse duration of 350 μs, which corresponds to a threshold pump power of about 230 W. Just above threshold, a considerably diminished maser burst, delayed in time and somewhat shorter in duration, could be observed.

We have constructed a simple model for how both the spin populations in the X and Z sublevels and the energization of (that is, photon number in) the maser mode change as functions of time in response to the applied optical pump pulse (Methods). The maser output power predicted by this model was fitted to the experimentally observed maser burst, yielding the green curve in Fig. 3, together with estimates of three spin-dynamics parameters. Our modelling suggests that near-zero-field operation is particularly advantageous for maser action (masing): the rate of spin–lattice relaxation is low and the transition linewidth is narrow[27].
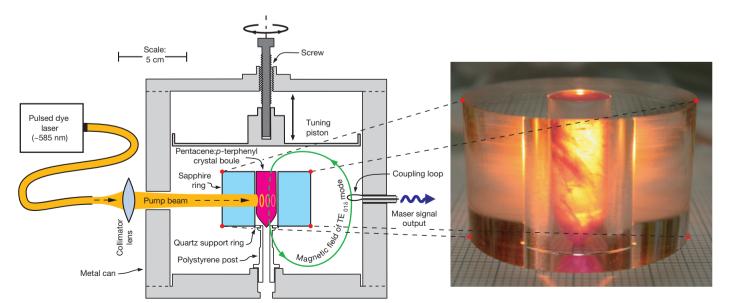

Figure 2 | Anatomy of the maser. A crystal of p-terphenyl doped with pentacene is located in the a.c. magnetic field of the $TE_{01\delta}$ mode of a microwave resonator and illuminated with a beam of yellow light from a pulsed dye laser. By maser action, the $TE_{01\delta}$ mode is energized, and from this mode a signal is extracted using a magnetic coupling loop. The photograph on the right shows the resonator's sapphire ring and the pentacene:p-terphenyl crystal slotted inside it; yellow back lighting reveals flaws within the crystal.
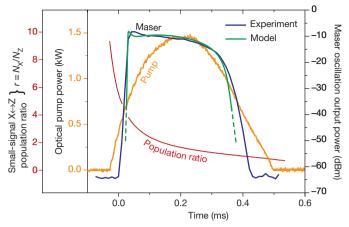
Figure 3 | Maser response in the time domain. The yellow trace (scale on near left) is the pump laser's instantaneous output power, which corresponds to an energy of ~0.5 J per pulse. The blue trace (scale on right) is the instantaneous power of the maser's oscillation burst as monitored with a microwave receiver at a resolution bandwidth of 10 MHz centred on 1.45 GHz. The green trace (same scale on right) is our fitted model's prediction of this instantaneous power. The red curve (scale on far left) is the predicted population ratio between the uppermost (X) and lowest (Z) maser levels in response to the optical pump in the low-signal regime.

Our experimental maser oscillator (Fig. 2) can be reconfigured as a pulsed-mode amplifier of weak electromagnetic signals by connecting its coupling loop to one port of a circulator[5]. A detailed, quantitative analysis of this amplifier's noise performance, taking into account deleterious sources of noise associated with losses within this circulator and the maser cavity itself, is provided in Methods. The red curve in Fig. 3 shows the population ratio $r \equiv N_X/N_Z$ between the uppermost (X) and lowest (Z) maser levels as predicted by our fitted model in response to the optical pump pulse shown (yellow trace). Whereas the blue and green curves correspond to saturation-limited maser oscillation, this red curve corresponds to the low-signal regime, where the energization of the maser mode is miniscule. Although decaying with time, r remains in excess of 2 for up to 150 μs from the onset of the pump pulse, implying that the intrinsic limit on the maser amplifier's residual noise temperature stays below 140 mK for the same amount of time. Whether or not r can be sustained sufficiently above 1 to allow continuous maser operation (given a suitable pump source and adequate cooling) would seem to be finely balanced with respect to the values of the relevant relaxation parameters.
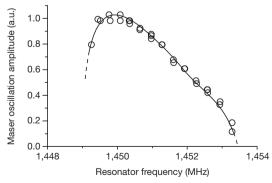
Figure 4 | Frequency response of maser action. The circles and associated guide line plot the measured plateau amplitude of the maser oscillation (vertical scale, obtained by down-converting the oscillation's output to a few tens of kilohertz and observing the resultant beat on a digital storage oscilloscope; a.u., arbitrary units) on illuminating the pentacene:p-terphenyl crystal (within its sapphire ring) with a ~0.5-J pulse of yellow light, for different frequencies of the microwave resonator's $TE_{01\delta}$ mode (horizontal scale).

When down-converted, sampled and Fourier-transformed, each burst of maser oscillation had a linewidth of ~6 kHz, consistent with that of a perfect sine wave modulated by the maser burst's envelope (Fig. 3, blue and green curves). Precise measurements of the maser's noise will be done if and when continuous operation is achieved.

With this device, we have demonstrated zero-field maser operation at room temperature with a solid-state gain medium. By improving the quality of the maser crystal, the effectiveness of the optical pumping and the Purcell factor of the resonator, we believe that the maser's threshold pump power can be substantially reduced, its efficiency (and, hence, output power as an oscillator) further increased and its residual noise temperature as an amplifier reduced. The application of a d.c. magnetic field bias should allow the maser's operating frequency to be tuned somewhat. The prospect of continuous maser operation, although not resolved here, remains a possibility and motivates attempts to identify different (yet perhaps structurally related) molecular systems whose properties are yet more advantageous for masing than those of pentacene:p-terphenyl.

Extremely low-noise amplifiers allow the detection of weak yet potentially important electromagnetic signals, which would otherwise go unnoticed. Low-phase-noise oscillators are essential components within frequency standards and clocks. Ultrasensitive parametric probes and detectors can, moreover, be derived from such oscillators. Because it can be operated at room temperature and in weak ambient magnetic fields (such as Earth's), the maser reported here thus opens new routes to the making of measurements at unprecedented levels of sensitivity and precision in areas such as space exploration, molecular biology and quantum computing. Furthermore, the general exploitation of intersystem crossing as an effective source of spin-polarized 'fuel' for the powering of more intricately orchestrated quantum molecular processes (beyond merely that of masing through a common microwave field) would seem to be a most fertile area of study.

METHODS SUMMARY

The maser crystal was grown using the vertical Bridgman–Stockbarger method in a double-walled glass vial[28]. Its constituent p-terphenyl was moderately zone-refined before growth; the substitutional concentration of its pentacene dopant was approximately 0.01%. The as-grown crystal boule was trimmed and then polished. It had extensive twinning (as revealed by spatially inhomogeneous birefringence) and other internal flaws (Fig. 2).

The microwave resonator was designed using both mode-matching and finite-element-method[29] software. The optical (c) axis of its monocrystalline sapphire ring lay parallel to the ring's geometric axis of rotation to within a tolerance of ±1°. All of the ring's surfaces and the internal surfaces of the resonator's silver-plated metal can were polished. On inserting the pentacene:p-terphenyl boule into the ring, the X↔Z transition was initially located by directly monitoring on an oscilloscope the response of an L-band Robinson oscillator bridge (which incorporated the resonator within its microwave circuit) to pulses of below-threshold yellow pump light.

The spin populations in the triplet ground state's X and Z sublevels, together with the photon number in the 1.45-GHz $TE_{01\delta}$ microwave mode, were modelled by coupled linear rate equations capturing the processes of stimulated emission and absorption, spin–lattice relaxation, and intersystem-crossing-mediated population into, and decay out of, the triplet sublevels[30]; see Methods for further details, including estimates of parameters that enter these equations as coefficients.

The predicted noise performance of our maser is based on formulae derived in ref. 5, which have previously been applied without low-temperature approximations to cryogenic ruby masers[6]. Noise contributions associated with losses in passive microwave components (including the resonator), as well as the 'intrinsic' maser noise associated with the effective spin temperature across the X↔Z transition, are all included in our analysis.

Full Methods and any associated references are available in the online version of the paper.

1. Schawlow, A. L. & Townes, C. H. Infrared and optical masers. *Phys. Rev.* **112**, 1940–1949 (1958).

2. Maiman, T. H. Stimulated optical radiation in ruby. *Nature* **187**, 493–494 (1960).
3. Kleppner, D., Goldenberg, H. M. & Ramsey, N. F. Properties of the hydrogen maser. *Appl. Opt.* **1**, 55–60 (1962).
4. Konoplev, I. V. *et al.* Experimental study of coaxial free-electron maser based on two-dimensional distributed feedback. *Phys. Rev. Lett.* **96**, 035002 (2006).
5. Siegman, A. E. *Microwave Solid-State Masers* (McGraw-Hill, 1964).
6. Clauss, R. C. & Shell, J. S. Ruby masers. In *Low-Noise Systems in the Deep Space Network* (ed. Reid, M. S.) (Deep Space Communication and Navigation Series, Jet Propulsion Laboratory, Caltech, 2008).
7. Dick, G. J. & Wang, R. T. Ultra-stable performance of the superconducting cavity maser. *IEEE Trans. Instrum. Meas.* **40**, 174–177 (1991).
8. Bourgeois, P.-Y. *et al.* Maser oscillation in a whispering-gallery-mode microwave resonator. *Appl. Phys. Lett.* **87**, 224104 (2005).
9. Mollier, J. C., Hardin, J. & Uebersfeld, J. Theoretical and experimental sensitivities of ESR spectrometers using maser techniques. *Rev. Sci. Instrum.* **44**, 1763–1771 (1973).
10. van der Waals, J. H. EPR of photo-excited triplet states: a personal account. *Appl. Magn. Reson.* **20**, 545–561 (2001).
11. Sloop, D. J., Yu, H.-L., Lin, T.-S. & Weissman, S. I. Electron spin echoes of a photoexcited triplet: pentacene in *p*-terphenyl crystals. *J. Chem. Phys.* **75**, 3746–3757 (1981).
12. Köhler, J., Brouwer, A. C. J., Groenen, E. J. J. & Schmidt, J. On the intersystem crossing of pentacene in *p*-terphenyl. *Chem. Phys. Lett.* **250**, 137–144 (1996).
13. Bogle, G. S. & Symmons, H. F. Zero-field masers. *Aust. J. Phys.* **12**, 1–20 (1959).
14. Whelehan, J. J. Low-noise amplifiers—then and now. *IEEE Trans. Microw. Theory Tech.* **50**, 806–813 (2002).
15. Wadefalk, N. *et al.* Cryogenic wide-band ultra-low-noise IF amplifiers operating at ultra-low DC power. *IEEE Trans. Microw. Theory Tech.* **51**, 1705–1711 (2003).
16. Smith, B. A. *et al.* A new look at the Saturn system: the Voyager 2 images. *Science* **215**, 504–537 (1982).
17. Asztalos, S. J. *et al.* SQUID-based microwave cavity search for dark-matter axions. *Phys. Rev. Lett.* **104**, 041301 (2010).
18. Korolev, A. M., Shnyrkov, V. I. & Shulga, V. M. Ultra-high frequency ultra-low dc power consumption HEMT amplifier for quantum measurements in millikelvin temperature range. *Rev. Sci. Instrum.* **82**, 016101 (2011).
19. Blank, A., Kastner, R. & Levanon, H. Exploring new active materials for low-noise room-temperature microwave amplifiers and other devices. *IEEE Trans. Microw. Theory Tech.* **46**, 2137–2144 (1998).
20. Hsu, H. & Tittel, F. K. Optical pumping of microwave masers. *Proc. IEEE* **51**, 185–189 (1963).
21. Sabisky, E. S. & Anderson, C. H. A solid-state cw optically pumped microwave maser. *Appl. Phys. Lett.* **8**, 298–300 (1966).
22. Blank, A. & Levanon, H. Toward maser action at room temperature by triplet-radical interaction and its application to microwave technology. *RIKEN Rev.* **44**, 128–130 (2002).
23. Blank, A. & Levanon, H. Applications of photoinduced electron spin polarization at room temperature to microwave technology. *Appl. Phys. Lett.* **79**, 1694–1696 (2001).
24. Antipov, S. *et al.* in *Proc. 23rd Particle Accelerator Conf.* 4535–4537, https://accelconf.web.cern.ch/accelconf/PAC2009/papers/fr5rfp001.pdf (Joint Accelerator Conference Website, 2009).
25. Kagawa, A., Murokawa, Y., Takeda, K. & Kitagawa, M. Optimization of ^{1}H spin density for dynamic nuclear polarization using photo-excited triplet electron spins. *J. Magn. Reson.* **197**, 9–13 (2009).
26. Purcell, E. M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **69**, 681–686 (1946).
27. Lang, J., Sloop, D. J. & Lin, T.-S. Orientational anisotropic studies by field rotation technique: near zero-field pulsed EPR experiments of pentacene in *p*-terphenyl. *J. Magn. Reson.* **176**, 249–256 (2005).
28. Selvakumar, S. *et al.* Growth and studies on SSVBT grown *p*-terphenyl single crystals. *J. Cryst. Growth* **275**, e265–e271 (2005).
29. Oxborrow, M. Traceable 2-D finite-element simulation of the whispering-gallery modes of axisymmetric electromagnetic resonators. *IEEE Trans. Microw. Theory Tech.* **55**, 1209–1218 (2007).
30. Takeda, K., Takegoshi, K. & Terao, T. Zero-field electron spin resonance and theoretical studies of light penetration into single crystals and polycrystalline material doped with molecules photoexcitable to the triplet state via intersystem crossing. *J. Chem. Phys.* **117**, 4940–4946 (2002).

Author Contributions J.D.B. performed initial pulsed electron paramagnetic resonance experiments with M.O., and analysed the resultant data. Both J.D.B. and M.O. carried out independent electromagnetic modelling of the TE$_{01\delta}$-mode microwave resonator using different software. M.O. grew the pentacene:*p*-terphenyl crystal, designed and constructed the microwave circuitry and optical systems, performed the final experiments and analysed the resultant data. N.M.A. initiated the original work on high-*Q* cavities. N.M.A. and M.O. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.O. (mo@npl.co.uk) or N.M.A. (n.alford@imperial.ac.uk).

METHODS

Maser crystal. The pentacene was of commercial origin and was donated by Sandrine Heutz of the Department of Materials, Imperial College London. The *p*-terphenyl was purchased from Alfa Aesar (99+% grade) and then moderately zone-refined, that is, sealed in glass tubes under dry argon and swept repeatedly at 0.5 mm min^{-1} up through a vertical furnace. A double-walled[28] vial of borosilicate glass (internal diameter, 18 mm; height, 70 mm; made by Cambridge Glassblowing Ltd) was charged with a mixture of pentacene and zone-refined *p*-terphenyl in a stoichiometric ratio of approximately 1:10^4. Owing to oxygen contamination during growth, the actual concentration of pentacene surviving in the crystal may have been somewhat lower than this. The vial was lowered through a vertical tube furnace at a rate of 1 mm h^{-1}. Although a clean, pink crystal formed initially, on cooling to room temperature at a rate of approximately 50 °C h^{-1}, many flaws appeared within it due to twinning and/or thermal shock. The vial was carefully cracked open and the crystal boule removed intact from it. This boule was then gently cut ('topped and tailed') with a circular saw and then polished, the last abrasive being opaline. The *p*-terphenyl crystal's *a*–*b* cleavage plane contained the boule's rotation axis to within ±5°. Extensive twinning meant that the crystal's *a* and *b* axes were differently oriented in different parts of the crystal (the pump beam always illuminated several such domains).

Microwave resonator. The resonator's basic aspect ratios and dimensions were modelled using in-house-written mode-matching software, with the aim of realizing a high-*Q* mode at 1.45 GHz. A '2.5-dimensional' FEM-based electromagnetic mode simulator[29] running in COMSOL, in conjunction with AutoCAD, was additionally used to determine suitable shapes and dimensions for all components. The polished sapphire ring was sourced from SurfaceNet GmbH. The metal can (including piston) was plated with a 30-μm layer of non-bright silver at Platerite (Aldershot, UK) and then polished by hand.

Modelling of the spin dynamics. We assume that the spin populations in the triplet ground state's X and Z sublevels, and the photon number in the 1.45-GHz TE$_{01\delta}$ microwave mode, are governed by coupled linear rate equations, with terms capturing the processes of stimulated emission and absorption, spin–lattice relaxation, and intersystem crossing-mediated population into, and decay out of, the triplet sublevels. (The Y sublevel is not included in the model as this would have introduced several more free parameters.) We ignore both the spatial variation of the TE$_{01\delta}$ mode's magnetic field across the pumped region of the pentacene:*p*-terphenyl crystal and the vectorial complexities associated with differently oriented pentacene molecules at different *p*-terphenyl lattice sites, in relation to the directions of both the (at least initially unpolarized) pump beam and the TE$_{01\delta}$ mode's magnetic field. Elaborations necessary to include the effects of inhomogeneous broadening and spin diffusion are not included; we assume a simple Lorentz-broadened X↔Z transition. We estimate that, owing to the large diameter of the pump beam and the low concentration of pentacene within the maser crystal (so increasing its absorption depth to the pump light), the heat generated by a single pump pulse shifts the frequency of the X↔Z transition by only a small fraction of its linewidth; we thus ignore thermal effects.

We have estimated (in part on the basis of a measurement using an optical power meter placed inside the sapphire ring) that, owing to reflections and scattering off poorly polished surfaces, only about 60% of the pump laser's output gets absorbed by the maser crystal; this optical loss is included in our model. On further assuming that the value for the room-temperature intersystem crossing yield is 62.5% (ref. 30) and that the values of the intersystem crossing population ratios ($p_X^{ISC} = 0.76, p_Z^{ISC} = 0.08$) and the decay lifetime of the Z sublevel ($\tau_Z = 10^{-3}$ s) are those determined in ref. 11 from X-band spin-echo measurements, we estimate that the ratio of the decay lifetimes between the Z and X sublevels is $\tau_Z/\tau_X \approx 8.5$,

that the effective linewidth of the X↔Z transition is $\Delta f \equiv 1/\pi T_2 \approx 860$ kHz and that the spin–lattice relaxation time is $T_1 \approx 135$ μs, with uncertainties of ±20% in all three parameters. These were the only free parameters fitted.

The green curve in Fig. 3 is the model prediction for the above parameter values. Our estimate of the linewidth agrees with the value of 700 kHz reported in ref. 27, and is broadly consistent with an estimate for the spin–spin relaxation time of $T_2 \approx 260$ ns extracted from a preliminary measurement on a small, more concentrated polycrystalline sample of pentacene:*p*-terphenyl using a standard X-band pulsed electron paramagnetic resonance spectrometer. Figure 4 would suggest, however, that our imperfect maser crystal suffers additional inhomogeneous broadening of a few megahertz. Our estimates of τ_Z/τ_X and T_1 are factors of two lower and several higher, respectively, when compared with the values of their high-magnetic-field counterparts, as determined in ref. 11. We have yet to study the zero-field Y↔Z and X↔Y transitions at ~1.34 GHz and ~107 MHz, respectively (the second requires a different sort of resonator suitable for VHF), as would be needed to determine all of the parameters within a complete model of the spin dynamics, including transitions to and from the Y sublevel.

Amplifier residual noise temperature. Configured as a cavity amplifier operated in reflection through a circulator, the maser's residual noise temperature can be formulated[5,6] as

$$T_{noise} = (1 + 1/G)\frac{L_{circ.}^{dB}}{e^{dB}} T_0 + (1 - 1/G)\left[\frac{L_{cav.}^{dB}}{G^{dB}} T_0 + \frac{G_{mas.}^{dB}}{G^{dB}} \frac{r}{r-1} \frac{hf_{mas.}}{k}\right] \quad (1)$$

where $X^{dB} \equiv 10 \log_{10}(X)$ denotes a quantity's decibel equivalent ($e^{dB} \approx 4.343$); $L_{circ.}^{dB}$ is the insertion loss through the circulator; $G^{dB} \equiv G_{mas.}^{dB} - L_{cav.}^{dB}$ is the amplifier's overall logarithmic power gain, where $G_{mas.}^{dB}$ is the gain associated with the masing process and $L_{cav.}^{dB}$ is the passive loss of the 'dark' (unpumped) cavity in reflection; T_0 is the operational temperature of the cavity and circulator, taken here to be 293 K; $f_{mas.} \approx 1.45$ GHz; and $r \equiv N_X/N_Z$ is the spin population ratio between the uppermost (X) and lowest (Z) maser levels. We now assume that the amplifier is good; that is, that its power gain $G \gg 1$ and that $G_{mas.}^{dB} \gg L_{cav.}^{dB}$. By using a low-loss circulator of say $L_{cav.}^{dB} \approx 0.1$ dB, the first term in equation (1) can be reduced to ~7 K. The middle term (proportional to $L_{cav.}^{dB}/G^{dB}$) originates in ohmic and/or dielectric losses in the cavity. When set to its maximum output power, the pump laser used in our experimental maser can provide sufficient light to exceed the masing threshold by a factor of approximately 50, that is, $L_{cav.}^{dB}/G^{dB} \approx 0.02$, giving the middle term a value of around 6 K; this noise contribution could be reduced further by devising a cavity with a higher magnetic Purcell factor or using a laser that supplies pulses of higher instantaneous pump power (though probably of only shorter duration).

This leaves the last, 'intrinsic' term in equation (1), which is proportional to $r/(r-1)$. The advantage of pumping optically via intersystem crossing is that r can be considerably greater than 1 even at room temperature. Naively taking this ratio to be that of the intersystem crossing process (Fig. 1) alone, that is, $r = 0.76/0.08 = 9.5$, the intrinsic term contributes a noise temperature of just 78 mK. In reality, r will be reduced by relaxation processes as time increases (Fig. 3, red curve), so raising this noise temperature. But the maser amplifier's overall noise temperature may still be deeply cryogenic for long enough to render the amplifier advantageous in certain pulse-mode applications such as the amplification of radar echoes. Various improvements are possible here, such as the construction of a travelling-wave maser[5,6], which would obviate the circulator's up-front noise contribution, in a narrow yet low-loss ('slow-wave') waveguide to keep $L_{cav.}^{dB}/G^{dB}$ low.

LETTER

Electronic read–out of a single nuclear spin using a molecular spin transistor

Romain Vincent[1], Svetlana Klyatskaya[2], Mario Ruben[2,3], Wolfgang Wernsdorfer[1] & Franck Balestro[1]

Quantum control of individual spins in condensed-matter devices is an emerging field with a wide range of applications, from nanospintronics[1,2] to quantum computing[3]. The electron, possessing spin and orbital degrees of freedom, is conventionally used as the carrier of quantum information in proposed devices[4,5,6,7,8,9]. However, electrons couple strongly to the environment, and so have very short relaxation and coherence times. It is therefore extremely difficult to achieve quantum coherence and stable entanglement of electron spins. Alternative concepts propose nuclear spins as the building blocks for quantum computing[10], because such spins are extremely well isolated from the environment and less prone to decoherence. However, weak coupling comes at a price: it remains challenging to address and manipulate individual nuclear spins[11,12,13,14]. Here we show that the nuclear spin of an individual metal atom embedded in a single-molecule magnet can be read out electronically. The observed long lifetimes (tens of seconds) and relaxation characteristics of nuclear spin at the single-atom scale open the way to a completely new world of devices in which quantum logic may be implemented.

Single-shot read-out of a single nuclear spin has remained elusive until the past few years, when it was achieved on a single nitrogen-vacancy defect centre in diamond[15]. The use of single-molecule magnets (SMMs) has also been proposed as a structural basis for nanospintronics[16] and quantum information[9]. The argument in favour of molecular building blocks for quantum computing is based on the possibility of synthesizing billions of perfectly identical SMMs, which can be integrated in devices by bottom-up approaches and for which the intrinsic magnetic properties and the environment can be chemically engineered. Moreover, experiments on an assembly of SMMs have already shown that: SMMs are characterized by a large

spin ground state combined with uniaxial magnetic anisotropy leading to two stable spin orientations; quantum phenomena such as quantum tunnelling of the magnetization (QTM)[17,18] and quantum phase interference[19] can be observed; and spin-state oscillations can be made coherent[20,21] for a time on the order of 1 μs. To achieve a longer coherence time, experiments must be done on the single-molecule level, avoiding interactions between SMMs. However, it is still difficult to detect single spin states in an SMM. Towards this goal, transport measurements have been made using scanning tunnelling microscopy (STM)[22,23], and SMMs have been studied in a transistor-like configuration using electromigrated junctions[24,25] or carbon-nanotube transistors[26]. So far, only signatures accounting for the electronic magnetic moment have been addressed, and there has been no quantitative comparison with the expected theoretical magnetic behaviour of an individual SMM.

Here, transport measurements taken through a single bis(phthalocyaninato)terbium(III) SMM (TbPc$_2$, Fig. 1) are studied in a three-terminal geometry obtained by electromigration[27]. TbPc$_2$ is a lanthanide SMM, in which the magnetic moment is carried by a single Tb^{3+} metal ion sandwiched between two organic phthalocyanine (Pc) ligands. We used this particular SMM for the several reasons. It has been reported that TbPc$_2$ SMMs conserve both their structural integrity and their magnetic properties even when sublimated at 820 K on a copper surface[28]. The redox state of the Tb^{3+} ion is very stable; that is, it is very unlikely that a current can flow through the Tb^{3+} ion (see Supplementary Information). The two Pc ligands have a conjugated π system, which can easily conduct electrons. Owing to strong spin–orbit coupling and the Pc ligand fields, TbPc$_2$ has a magnetic ground state of $J = 6$ and a strong uniaxial magnetic anisotropy (Ising-like, Fig. 1a). The ground-state doublet $J_z = \pm 6$ is well isolated
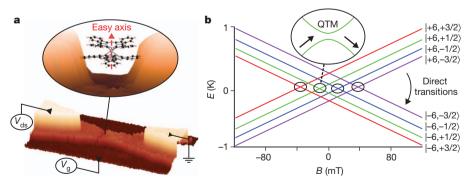
Figure 1 | Geometry of the molecular spin transistor and magnetization reversal processes. a, Three-dimensional extrapolation of a scanning-electron-microscope image showing the most favourable structure of the single-molecule-based transistor. A schematic zoom into the nano gap shows the molecular structure of the TbPc$_2$ SMM and its easy axis. The charge state of the ligand read-out dot can be controlled by the gate voltage, V_g, and the voltage difference between the electrodes is controlled through the drain-source voltage, V_{ds}. **b,** Zeeman diagram presenting the energy, E, of the two ground

states, $J_z = \pm 6$, as a function of the magnetic field (B). The two ground states are each split into four different sub-states owing to the hyperfine coupling with the nuclear spin of $I = 3/2$. Coloured lines denote the I_z components of the nuclear spin states: purple, $-3/2$; blue, $-1/2$; green, $1/2$; and red, $3/2$. Two processes are responsible for the magnetization reversal. In small magnetic fields, QTM can occur at the avoided energy-level crossings with the same I_z but opposite J_z, indicated by the black circles. In higher fields, a direct relaxation process can lead to the reversal of J.

[1]Institut Néel, CNRS et Université Joseph Fourier, BP 166, F-38042 Grenoble Cedex 9, France. [2]Institute of Nanotechnology, Karlsruhe Institute of Technology, 76344 Eggenstein-Leopoldshafen, Germany. [3]Institut de Physique et Chimie des Matériaux de Strasbourg, CNRS-Université de Strasbourg, 67034 Strasbourg, France.

from the excited states by an energy gap of a few hundred kelvin[29]. The excited states can thus be neglected for experiments carried out at low temperatures (less than 5 K) and small magnetic fields (less than 10 T). The ligand fields also induce transverse magnetic anisotropy, which is responsible for the coupling of the two states of the doublet $J_z = \pm 6$, so that energy-level crossings are avoided[29] (see the black circles in the Zeeman diagram of Fig. 1b). Finally, because of strong hyperfine coupling with the nuclear spin $I = 3/2$ of the Tb^{3+} ion, each ground state $J_z = \pm 6$ is split into four states (Fig. 1b).

Within the low-temperature regime, the magnetization reversal of a $TbPc_2$ SMM takes place through two different processes. In small magnetic fields, QTM at the four avoided level crossings (see above) is the main mechanism[29], whereas in higher magnetic fields, relaxation processes can be direct, involving non-coherent tunnelling events combined with phonon emission. QTM of a $TbPc_2$ SMM has been observed in a supramolecular carbon-nanotube SMM spin-valve geometry[26], but the signature of the hyperfine coupling could not be accessed, owing to the background noise induced by the surrounding $TbPc_2$ SMMs. Here, we show that by detecting the reversal of the electronic magnetic moment of the $TbPc_2$ SMM, we can produce an electronic read-out of the four nuclear spin states of a single Tb^{3+} ion with high efficiency and in perfect quantitative agreement with the theoretical predictions[29]. Furthermore, we estimate the lifetimes and the temperature of a single nuclear spin.

Transport measurements through a single $TbPc_2$ SMM have been taken previously by scanning tunnelling spectroscopy (STS) experiments[22], for which electronic transport occurred through the Pc ligands and exhibited Coulomb blockade and Kondo effects depending on its charge state (spin states $S = 0$ or $1/2$). However, no signature of the magnetic moment carried by the Tb^{3+} ion was observed in this experiment. To detect the reversal of the magnetic moment, and to produce the electronic read-out of the Tb^{3+} nuclear spin states, we

inserted the $TbPc_2$ SMM directly into a gold junction obtained by the electromigration technique (see Fig. 1a and Methods). The differential conductance, dI/dV, as the function of the drain-source voltage, V_{ds}, and the gate voltage, V_g, is presented in Fig. 2a. It exhibits a single charge-degeneracy point with a weak spin $S = 1/2$ Kondo effect. A detailed study of the Kondo peak as a function of the applied magnetic field is presented in the Supplementary Information. A ferromagnetic exchange interaction of about 0.35 T was measured between the spin 1/2 of the quantum dot and the magnetic moment carried by the Tb^{3+} ion. Alternative coupling mechanisms (dipolar, magneto-Coulomb, mechanical, flux) are discussed in the Supplementary Information, but the relatively high value of the exchange interaction means that it should be the major contribution to the coupling mechanism, and that the quantum dot is spatially located close to the Tb^{3+} ion. This is indirect proof that the electronic transport occurs through the aromatic Pc ligands, and that the most favourable geometric configuration is the one depicted in Fig. 1a. To summarize: the Pc ligands form a molecular quantum dot and the anisotropic magnetic moment of the Tb^{3+} ion is coupled to the electron path only indirectly, mainly through a ferromagnetic exchange interaction; the presence of a gate allows fine-tuning from the Coulomb blockade to the Kondo regime of the molecular quantum dot; the magnetic properties of the Tb^{3+} ion are then independent of the charge state of the Pc quantum dot (see Supplementary Information); and owing to the exchange coupling, we can use the Pc ligands as a read-out quantum dot to detect the reversal of the electronic magnetic moment carried by the Tb^{3+} ion spin dot.

To produce the electronic read-out of the single nuclear spin carried by the spin dot, we chose experimental conditions close to the charge-degeneracy point ($V_g = -0.9$ V and $V_{ds} = 0$ V in Fig. 2a, see Supplementary Information). When sweeping the magnetic field from negative to positive values, we observed a single abrupt jump of the differential conductance, which reversed when we swept the field in the opposite direction (Fig. 2b). Our detailed studies (see below) showed that these jumps can be accounted for by the reversals of the Tb^{3+} magnetic moment, which slightly influence the chemical potential of the read-out quantum dot through the magnetic interactions
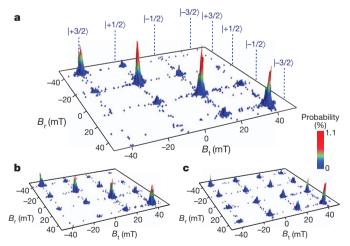

Figure 2 | Conductance characteristics and electronic read-out procedure. a, Stability diagram of the Pc read-out quantum dot exhibiting the differential conductance, dI/dV, in units of the quantum of conductance, G_0, as a function of gate voltage, V_g, and bias voltage, V_{ds}, at 0.1 K. **b**, dI/dV measurements for a given working point ($V_g = -0.9$ V; $V_{ds} = 0$ V) as function of the magnetic field, B. The arrows indicate the field-sweep direction. Abrupt jumps in the differential conductance, attributed to the switching of the Tb^{3+} magnetic moment, are visible for all traces of B, showing a clear hysteresis in the dI/dV characteristics. **c**, Histogram of switching field obtained for 11,000 field sweeps showing four preferential field values that are assigned to QTM events. **d**, Normalized hysteresis loop of a single $TbPc_2$ SMM obtained by integration of 1,000 field sweeps and performed for trace and retrace on a larger magnetic-field range than in **c**. The four arrows on the trace curve show the four preferential field values associated to QTM (red, −40 mT; green, −14 mT; blue, 14 mT; purple, 40 mT).

Figure 3 | Transition matrix of the QTM events as a function of the waiting time. The switching fields of the Tb^{3+} magnetic moment of subsequent field sweeps are plotted in two-dimensional histograms for three waiting times, t_w: **a**, $t_w = 0$ s; **b**, $t_w = 20$ s; and **c**, $t_w = 50$ s. The two axes correspond to the trace and retrace field sweeps, B_t and B_r, respectively. Two successive measurements with the same nuclear spin states are situated on the diagonal of the matrix, whereas the off-diagonal positions correspond to nuclear spin-state changes of $\Delta m_I = \pm 1, \pm 2$ and ± 3. The predominance of diagonal terms up to $t_w = 20$ s indicates the long level lifetime of the nuclear spin states. For $t_w = 50$ s, the diagonal terms vanish owing to nuclear spin-flip processes. Furthermore, the high amplitude of the bottom-right ($B_t = B_r = 40$ mT) matrix element accounts for the relaxation of the nuclear spin towards a thermal equilibrium.

Figure 4 | Spin-flip dynamics and nuclear spin-state occupancy of the Tb^{3+} nuclear spin states. Evolution of the nuclear spin-state occupancy as a function of the waiting time, t_w, for two different working points: **a**, gate voltage $V_g = -0.9$ V; and **b**, $V_g = -0.1$ V (both at bias voltage $V_{ds} = 0$ V). The measurements clearly show that the populations evolve towards different thermal equilibria. **c**, Spin dynamics for a fixed waiting time of 10 s, as a function of temperature, T. With increasing temperature, the population of the different spin states evolves towards equal occupancy.

(see Supplementary Information). We took subsequent measurements of the reversal as a function of the magnetic-field angle in a dilution refrigerator equipped with a home-made three-dimensional vector magnet. We were thus able to align the field-sweep direction accurately along the easy axis of the TbPc$_2$ SMM (see Supplementary Information).

We obtained the reversal statistics of the magnetization by sweeping the magnetic field back and forth at 0.05 T s^{-1} and measuring histograms of the switching fields. We present a typical example in Fig. 2c, showing four preferential magnetic-field values. These are in perfect quantitative agreement with theoretical predictions[29] of QTM of a Tb^{3+} magnetic moment at the four avoided energy-level crossings with the nuclear spin states $|3/2\rangle$, $|1/2\rangle$, $|-1/2\rangle$ and $|-3/2\rangle$ (Fig. 1b). Moreover, the histograms of the four switching fields do not overlap; this shows the high efficiency of this electronic read-out procedure. The field integration of the normalized switching histograms yields the magnetic hysteresis loop (Fig. 2d), which is in excellent accordance with micro-SQUID measurements of assemblies of TbPc$_2$ SMMs (see Supplementary Information and ref. 29). In higher magnetic fields, the reversal of the magnetization occurs stochastically, as predicted for a direct relaxation process involving a non-coherent tunnelling event combined with a phonon emission.

The lifetimes of the four nuclear spin states could be measured by studying the correlations between subsequent measurements as a function of the waiting time, t_w, between field sweeps. Figure 3 presents two-dimensional histograms (transition matrices) obtained from 22,000 field sweeps. The two axes correspond to the trace and retrace field sweeps, B_t and B_r. Two subsequent measurements with the same nuclear spin states are situated on the diagonal of the matrix; the off-diagonal positions correspond to nuclear spin state changes of $\Delta m_I = \pm 1$, ± 2, and ± 3. For zero waiting time ($t_w = 0$ s, Fig. 3a), the diagonal positions are predominant, highlighting the robustness of the nuclear spin states; that is, indicating long level lifetimes for the individual Tb^{3+} nuclear spin states. The diagonal positions persist even for a waiting time of $t_w = 20$ s (Fig. 3b). However, for $t_w = 50$ s, the off-diagonal positions start to be populated, which suggests the occurrence of nuclear spin–flip processes during the waiting time (Fig. 3c). From this series of measurements, we conclude that the level lifetime of the nuclear spin states is on the order of tens of seconds, confirming that the invasiveness of the measurement procedure is low (see Supplementary Information).

More detailed insight into the spin–flip dynamics of an individual nuclear spin can be gained by measuring the population of nuclear spin states as a function of waiting time and temperature (Fig. 4). To this end, we determined the nuclear spin-state occupancy for two different working points corresponding to two different charge states of the read-out quantum dot ($V_g = -0.9$ V in Fig. 4a and $V_g = -0.1$ V in Fig. 4b, both at $V_{ds} = 0$ V). It is clear that the state occupancy, relaxing towards a thermal equilibrium, depends heavily on the transport characteristics (the current flowing through the Pc quantum dot and/or the electrostatic environment modulated by the gate voltage). Indeed, the electron tunnelling through the Pc-ligands read-out dot gives rise to small fluctuations of the local electric field, which could modify nuclear spin–flip processes through the quadrupole interaction[30]. A detailed study addressing this problem is in progress, and beyond the scope of the work presented here. We also determined the population of nuclear spin states as a function of the temperature (Fig. 4c). The strong temperature dependence of this occupancy demonstrates that a single nuclear spin can be thermalized down to at least 0.2 K, which is close to the electronic temperature of our dilution refrigerator (0.08 K).

In conclusion, we have performed electronic transport studies on a TbPc$_2$ SMM in a transistor-like set-up, taking measurements of a single nuclear spin. Our experimental read-out procedure relies on the highly efficient detection of the QTM of the electronic magnetic moment at particular values of the magnetic field corresponding to the four avoided energy-level crossings; this is in perfect quantitative agreement with theoretical predictions. Using this procedure, the dynamics of the four different nuclear spin states of a single Tb^{3+} ion could be electronically determined. The observation of energy-level lifetimes on the order of tens of seconds opens the way to coherent manipulations of a single nuclear spin. Moreover, because of the different energy-level spacing that originates from the nuclear quadrupole interaction term (see Supplementary Information), this particular device is presumed to be an excellent candidate to perform the Grover algorithm. Indeed it has been proposed that the coherent manipulation of a nuclear spin $I = 3/2$ can be performed through a multiphoton process using this algorithm[31]. It was predicted that the greater the quadrupolar constant, the better the implementation of the algorithm would be. Note that the value presented in ref. 31 was three orders of magnitude smaller than the value presented here ($P = 0.013$ cm^{-1}). Thus, the demonstrated possibility of addressing and detecting single nuclear spin states using the QTM of SMMs, in combination with the observation of long energy-level lifetimes, opens up a bright new world of nanospintronics with integrated memory, logics and possibly quantum logics.

METHODS SUMMARY

The single-molecule transistor was prepared by blow-drying a dilute toluene solution of the TbPc$_2$ molecule onto a gold nanowire on an Au/HfO$_2$ gate fabricated through atomic-layer deposition. Before the solution was blow-dried, the electrodes were cleaned with acetone, ethanol, isopropanol solution and oxygen plasma. The connected sample was enclosed in a high-frequency, low-temperature filter, consisting of a thermocoax microwave filter and π filters, anchored to the mixing chamber of a dilution refrigerator with a base temperature of about 0.05 K. The molecule-coated nanowire was then broken by electromigration, using a voltage ramp at 4 K. The electromigration technique used real-time electronics to increase the yield of coupling a single molecule to the electrodes.

Transport measurements were taken using a lock-in amplifier in a dilution refrigerator with an electronic temperature of about 0.08 K. It was equipped with a home-made three-dimensional vector magnet, allowing us to sweep the magnetic field in three dimensions at rates up to 0.2 T s^{-1}.

1. Awschalom, D. D., Loss, D. & Samarth, N. *Semiconductor Spintronics and Quantum Computation* (Springer, 2002).

2. Zutic, I., Fabian, J. & Das Sarma, S. Spintronics: fundamentals and applications. *Rev. Mod. Phys.* **76,** 323–410 (2004).
3. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
4. Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57,** 120–126 (1998).
5. Imamoglu, A. *et al.* Quantum information processing using quantum dot spins and cavity QED. *Phys. Rev. Lett.* **83,** 4204–4207 (1999).
6. Eble, B. *et al.* Hole–nuclear spin interaction in quantum dots. *Phys. Rev. Lett.* **102,** 1–4 (2009).
7. Jelezko, F., Gaebel, T., Popa, I., Gruber, A. & Wrachtrup, J. Observation of coherent oscillations in a single electron spin. *Phys. Rev. Lett.* **92,** 1–4 (2004).
8. Vrijen, R. *et al.* Electron-spin-resonance transistors for quantum computing in silicon-germanium heterostructures. *Phys. Rev. A* **62,** 1–10 (2000).
9. Leuenberger, M. N. & Loss, D. Quantum computing in molecular magnets. *Nature* **410,** 789–793 (2001).
10. Kane, B. E. A silicon-based nuclear spin quantum computer. *Nature* **393,** 133–137 (1998).
11. Chuang, I. L., Gershenfeld, N. & Kubinec, M. Experimental implementation of fast quantum searching. *Nature* **393,** 143–146 (1998).
12. Chuang, I. L., Vandersypen, L. M. K., Zhou, X., Leung, D. W. & Lloyd, S. Experimental realization of a quantum algorithm. *Phys. Rev. Lett.* **80,** 3408–3411 (1998).
13. DiVincenzo, D. P. Two-bit gates are universal for quantum computation. *Phys. Rev. A.* **51,** 1015–1022 (1995).
14. Berman, G. P., Doolen, G. D., Hammel, P. C. & Tsifrinovich, V. I. Solid-state nuclear-spin quantum computer based on magnetic resonance force microscopy. *Phys. Rev. B* **61,** 14694–14699 (2000).
15. Neumann, P. *et al.* Single-shot readout of a single nuclear spin. *Science* **329,** 542–544 (2010).
16. Bogani, L. & Wernsdorfer, W. Molecular spintronics using single-molecule magnets. *Nature Mater.* **7,** 179–186 (2008).
17. Friedman, J. R., Sarachik, M. P., Tejada, J. & Ziolo, R. Macroscopic measurement of resonant magnetization tunnelling in high-spin molecules. *Phys. Rev. Lett.* **76,** 3830–3833 (1996).
18. Thomas, L. *et al.* Macroscopic quantum tunnelling of magnetization in a single crystal of nanomagnets. *Nature* **383,** 145–147 (1996).
19. Wernsdorfer, W. & Sessoli, R. Quantum phase interference and parity effects in magnetic molecular clusters. *Science* **284,** 133–135 (1999).
20. Ardavan, A. *et al.* Will spin-relaxation times in molecular magnets permit quantum information processing? *Phys. Rev. Lett.* **98,** 1–4 (2007).
21. Bertaina, S. *et al.* Quantum oscillations in a molecular magnet. *Nature* **453,** 203–206 (2008).
22. Komeda, T. *et al.* Observation and electric current control of a local spin in a single-molecule magnet. *Nature Commun.* **2,** 217 (2011).
23. Otte, A. F. *et al.* The role of magnetic anisotropy in the Kondo effect. *Nature Phys.* **4,** 847–850 (2008).
24. Heersche, H. B. *et al.* Electron transport through single Mn_{12} molecular magnets. *Phys. Rev. Lett.* **96,** 206801 (2006).
25. Zyazin, A. S. *et al.* Electric field controlled magnetic anisotropy in a single molecule. *Nano Lett.* **10,** 3307–3311 (2010).
26. Urdampilleta, M., Cleuziou, J.-P., Klyatskaya, S., Ruben, M. & Wernsdorfer, W. Supramolecular spin valves. *Nature Mater.* **10,** 502–506 (2011).
27. Park, H., Lim, A. K. L., Alivisatos, A. P., Park, J. & McEuen, P. L. Fabrication of metallic electrodes with nanometer separation by electromigration. *Appl. Phys. Lett.* **75,** 301–303 (1999).
28. Stepanow, S. *et al.* Spin and orbital magnetic moment anisotropies of monodispersed bis(phthalocyaninato)terbium on a copper surface. Supp. Inf. *J. Am. Chem. Soc.* **132,** 11900–11901 (2010).
29. Ishikawa, N., Sugita, M. & Wernsdorfer, W. Quantum tunneling of magnetization in lanthanide single-molecule magnets: bis(phthalocyaninato)terbium and bis(phthalocyaninato)dysprosium anions. *Angew. Chem. Int. Ed.* **44,** 2931–2935 (2005).
30. Abragam, A. *The Principles of Nuclear Magnetism* (Oxford Univ. Press, 1994).
31. Leuenberger, M. N. & Loss, D. The Grover algorithm with large nuclear spins in semiconductors. *Phys. Rev. B* **68,** 165317 (2003).

LETTER

Doubling of marine dinitrogen–fixation rates based on direct measurements

Tobias Großkopf[1]*, Wiebke Mohr[1]*†, Tina Baustian[1], Harald Schunck[1], Diana Gill[1], Marcel M. M. Kuypers[2], Gaute Lavik[2], Ruth A. Schmitz[3], Douglas W. R. Wallace[4] & Julie LaRoche[1]†

Biological dinitrogen fixation provides the largest input of nitrogen to the oceans, therefore exerting important control on the ocean's nitrogen inventory and primary productivity[1–3]. Nitrogen-isotope data from ocean sediments suggest that the marine-nitrogen inventory has been balanced for the past 3,000 years (ref. 4). Producing a balanced marine-nitrogen budget based on direct measurements has proved difficult, however, with nitrogen loss exceeding the gain from dinitrogen fixation by approximately 200 Tg N yr^{-1} (refs 5, 6). Here we present data from the Atlantic Ocean and show that the most widely used method of measuring oceanic N$_2$-fixation rates[7] underestimates the contribution of N$_2$-fixing microorganisms (diazotrophs) relative to a newly developed method[8]. Using molecular techniques to quantify the abundance of specific clades of diazotrophs in parallel with rates of ^{15}N$_2$ incorporation into particulate organic matter, we suggest that the difference between N$_2$-fixation rates measured with the established method[7] and those measured with the new method[8] can be related to the composition of the diazotrophic community. Our data show that in areas dominated by *Trichodesmium*, the established method underestimates N$_2$-fixation rates by an average of 62%. We also find that the newly developed method yields N$_2$-fixation rates more than six times higher than those from the established method when unicellular, symbiotic cyanobacteria and γ-proteobacteria dominate the diazotrophic community. On the basis of average areal rates measured over the Atlantic Ocean, we calculated basin-wide N$_2$-fixation rates of 14 ± 1 Tg N yr^{-1} and 24 ± 1 Tg N yr^{-1} for the established and new methods, respectively. If our findings can be extrapolated to other ocean basins, this suggests that the global marine N$_2$-fixation rate derived from direct measurements may increase from 103 ± 8 Tg N yr^{-1} to 177 ± 8 Tg N yr^{-1}, and that the contribution of N$_2$ fixers other than *Trichodesmium* is much more significant than was previously thought.

The oceans support roughly half of Earth's biological carbon fixation[9]. Carbon itself is rarely considered to be limiting in the ocean; instead, elements such as nitrogen, phosphorus and iron control primary productivity[2]. Of these elements, nitrogen is special because it is always present in large amounts as dissolved N$_2$, but this form is available only to the diazotrophs, a restricted group of prokaryotes[1]. Biological N$_2$ fixation is the largest source of fixed-nitrogen input into the ocean and it has been proposed that this controls marine primary productivity on geological timescales[10]. Whereas phytoplankton seem to be limited in growth over large areas by the availability of fixed N (for example, nitrate, nitrite or ammonium), the productivity of diazotrophs is controlled by other environmental factors, such as the availability of Fe and P (ref. 11). Thus, increased desertification owing to changes in land use may promote marine N$_2$ fixation through increased aeolian input of iron[12], whereas the increased use

of fertilizers and riverine run-off or increased atmospheric N deposition may have the opposite effect[13,14]. An understanding of marine N$_2$ fixation and its response to anthropogenic forcing is crucial for assessing the future of oceanic primary productivity. However, recent attempts to produce a balanced N budget have usually left a major gap on the input side[5,6]. This creates a paradox, because stable isotopes recorded in ocean sediments suggest that the marine N inventory has been balanced for the past 3,000 years (ref. 4). The paradox can be reconciled in three ways[6]. One hypothesis is that current, transient N-loss rates are exceeding N$_2$ fixation, but that over timescales of several thousand years, variable losses and gains act to maintain a steady-state fixed-N inventory[15,16]. Two alternate hypotheses invoke the possibilities that either the N-loss or the N-gain terms are in error. Several lines of evidence point out that even the most conservative estimates of oceanic N losses are higher than current estimates of N input based on field measurements of N$_2$ fixation. This suggests that contemporary N$_2$-fixation rates may have been underestimated grossly[5,6].

In the past decade, molecular techniques have led to the discovery of a variety of previously unrecognized diazotrophs that express their nitrogenase genes and fix N$_2$, thereby showing that there are gaps in our knowledge of marine N$_2$ fixation[17–21]. The discovery of N$_2$ fixation in mesopelagic waters in the Pacific[22,23] and the possibility of significant involvement of heterotrophic bacteria in global marine N$_2$ fixation[24,25] reinforce the idea that we are missing a large fraction of global marine N input by N$_2$ fixation. Biogeochemical modelling and geochemical-tracer techniques have been used to estimate regional and global magnitudes of marine N$_2$ fixation, but these indirect approaches rely on assumptions that require verification by direct biological rate measurements of N$_2$ fixation.

A recent laboratory study showed that the established method for measuring oceanic N$_2$ fixation (accounting for around 90% of the published rates[26]) leads to an underestimate, owing to slow and incomplete equilibration between the water sample and the ^{15}N$_2$ tracer added as a gas bubble[8]. If this finding holds for measurements of environmental N$_2$-fixation rates, an upward revision of the largest source of combined N into the ocean—that is, N$_2$-fixation—would be inevitable. We investigated the magnitude of the underestimation of N$_2$-fixation rates in open-ocean settings. In autumn 2009, on two research cruises to the Atlantic Ocean between 25° N and 45° S, we compared the established ^{15}N$_2$ tracer method for measuring N$_2$ fixation[7] (hereafter called the bubble-addition method) with a recently developed method[8] in which the ^{15}N$_2$ tracer is added as a dissolved gas (the dissolution method). Our sampling covered a wide variety of oceanic conditions with sea surface temperatures ranging from 10 °C to more than 28 °C (Fig. 1). At each station, we conducted parallel incubations with the dissolution and the bubble-addition methods, using dual labelling with NaH^{13}CO$_3$ and ^{15}N$_2$ gas in the same incubation bottles to measure

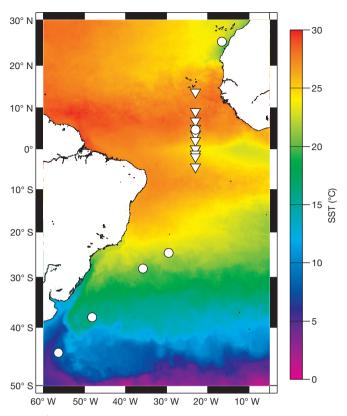**Figure 1 | Sampling sites and sea surface temperature.** Map of average night-time sea surface temperature (SST) in °C in autumn 2009 in the Atlantic Ocean (http://oceancolor.gsfc.nasa.gov/), overlaid with stations sampled on the RV *Meteor* cruise 80–81 (triangles) and the RV *Polarstern* cruise ANT-XXVI/1 (circles). Triangles also correspond to 11 vertical profile stations used in Fig. 3. The equatorial station was sampled twice during the course of the cruise, so two triangles are superimposed at this location.

dissolved inorganic carbon (DIC) assimilation and N_2 fixation simultaneously with both methods (see Supplementary Information for experimental details).

The ^{13}C-labelled DIC was added to check for systematic differences in biological activity attributable to the differences in methods, and the results showed that carbon-fixation rates were in good agreement between the two experimental methods (Fig. 2b). By contrast, there was a large difference and poor overall correlation between the N_2-fixation rates calculated using the two methods (Fig. 2a). Average depth-integrated rates of N_2 fixation over the whole Atlantic Ocean (25° N to 45° S) differed by a factor of 1.7 (91 ± 4

and $54 \pm 2\,\mu\text{mol N m}^{-2}\,\text{d}^{-1}$ for the dissolution and bubble-addition methods, respectively; mean and propagated s.e.m., $n = 17$; Wilcoxon matched pairs test $P < 0.01$). Furthermore, a geographical pattern emerged among the differences in the rates measured with the two methods. In the northern part of the 23° W section (15° N to 5° S), the depth-integrated N_2-fixation rates derived from four vertical profiles were, on average, 62% higher with the dissolution method than with the bubble-addition method (194 ± 12 and $120 \pm 5\,\mu\text{mol N m}^{-2}\,\text{d}^{-1}$, respectively; means and propagated s.e.m., $n = 4$). In the equatorial Atlantic Ocean (4.5° N to 5° S), a region previously not considered important for N_2 fixation[27], the depth-integrated N_2-fixation rates were on average 570% higher with the dissolution method ($55 \pm 7\,\mu\text{mol N m}^{-2}\,\text{d}^{-1}$) than with the bubble-addition method ($8 \pm 1\,\mu\text{mol N m}^{-2}\,\text{d}^{-1}$; means and propagated s.e.m., $n = 7$) (Fig. 3). Moreover, in the South Atlantic Gyre (38° S) and in the Falkland Current (44° S), where water temperatures were as low as 16° C and 10° C, respectively, considerable N_2 fixation was detected with the dissolution method (0.44 ± 0.1 and $0.54 \pm 0.1\,\mu\text{mol N m}^{-3}\,\text{d}^{-1}$; means and s.e.m., $n = 3$). The rates detected with the bubble-addition method were substantially lower (0.10 ± 0.01 and $0.18 \pm 0.03\,\mu\text{mol N m}^{-3}\,\text{d}^{-1}$ for 38° S and 44° S, respectively; means and s.e.m., $n = 3$).

We used TaqMan assays based on the presence of the nitrogenase reductase (*nifH*) gene to quantify the relative abundance of the diazotrophic phylotypes found in the tropical Atlantic Ocean[20,21]. We identified two geographically separated areas, dominated by distinct diazotrophic phylotypes (Fig. 4 and Supplementary Fig. 2). The most abundant diazotroph was *Trichodesmium*, peaking at 9° N with 5×10^5 *Trichodesmium nifH* copies per litre. We measured the highest N_2-fixation rates inside a *Trichodesmium* bloom at 13.75° N with $360 \pm 8.5\,\mu\text{mol N m}^{-2}\,\text{d}^{-1}$ and $219 \pm 8.2\,\mu\text{mol N m}^{-2}\,\text{d}^{-1}$ measured with the dissolution and bubble-addition methods, respectively (means and s.e.m., $n = 3$). Although *Trichodesmium* was present throughout the whole area (Fig. 4), its abundance declined rapidly south of 5° N. Diatom-associated heterocystous diazotrophs, unicellular cyanobacteria and γ-proteobacteria showed a peak of abundance within the *Trichodesmium* bloom and a second peak in the equatorial Atlantic region, where they outnumbered *Trichodesmium*. We characterize the tropical North Atlantic (5° N to 15° N) as a region of *Trichodesmium* dominance, where underestimation of N_2-fixation rates by the bubble-addition method was less severe (but still significant). By contrast, we found the largest underestimation (570%) with the bubble-addition method in the equatorial Atlantic (4.5° N to 5° S), which was dominated by diazotrophs other than *Trichodesmium* (Supplementary Figs 4 and 5). The combined results indicate that the magnitude of underestimation in N_2-fixation rates measured with the bubble-addition method relative to the
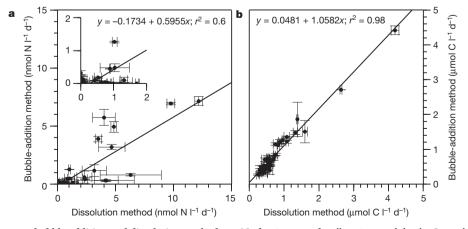

Figure 2 | Comparison between bubble-addition and dissolution methods. a, N_2-fixation rates for all stations and depths. Inset shows zoom on the 0–2-nmol l^{-1} d^{-1} range. **b**, Carbon-fixation rates for all stations and depths. Error bars indicate standard errors of triplicate incubations (s.e.m., $n = 3$).
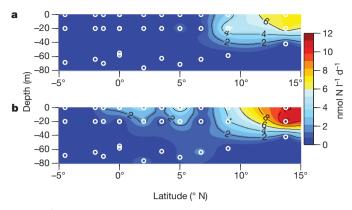
Figure 3 | Mixed-layer inventory of N₂-fixation rates in the tropical and equatorial Atlantic Ocean. Stations correspond to triangles in Fig. 1. **a**, N₂-fixation rates measured with the bubble-addition method. **b**, N₂-fixation rates measured with the dissolution method.

dissolution method was related to the composition of the diazotrophic community (Supplementary Fig. 2).

The species composition of the diazotrophic community can affect the level of underestimation for several reasons. In the bubble-addition method, the equilibration time for a $^{15}N_2$ gas bubble is longer than the typical incubation time of 24 h, despite manual shaking at the onset of the incubation[8] (Supplementary Fig. 6). This means that the bubble and its immediate surroundings will remain highly enriched in $^{15}N_2$, whereas water at the bottom of the incubation bottle will be less enriched (see Supplementary Information). As a result, buoyant diazotrophs (such as *Trichodesmium*) may be exposed to a higher fraction of the added $^{15}N_2$ label, whereas diazotrophs that sink (such as *Richelia*, when in symbiotic relationship with diatoms[28]) and accumulate at the bottom of the bottle may be exposed to a lower fraction of the $^{15}N_2$. Underestimation bias will also be greater when
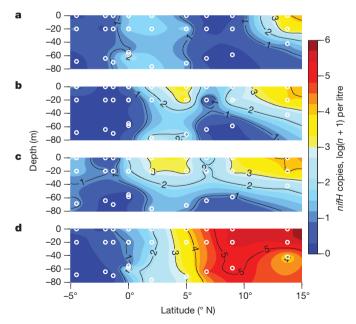
Figure 4 | Relative abundance of various phylotypes of diazotrophic bacteria from the same stations as the N₂-fixation rate measurements, estimated with TaqMan *nifH* gene assays. **a**, Unicellular cyanobacteria (Group A, Group B and Group C cyanobacteria). **b**, Diatom-associated heterocystous cyanobacteria. **c**, Diazotrophic γ-proteobacteria. **d**, *Trichodesmium*. There was no DNA sampling at the 3.5° N station. Note that the *nifH* gene copies (*n*) per litre of diatom-associated heterocystous cyanobacteria (such as *Richelia*) detects all cells of the symbionts, not just the heterocysts in which N₂ fixation is actively taking place.

the start of incubation coincides with the peak of N₂-fixation activity of the dominant diazotrophs (such as *Crocosphaera* for night-time incubations and *Trichodesmium* for daytime ones). These biases are avoided with the dissolution method because the $^{15}N_2$ label is added as a gas dissolved in sea water that is mixed uniformly with the sample at the start of the incubation.

Using published measurements of N₂-fixation rates obtained with the bubble-addition method, we calculated basin-wide and global budgets of marine N₂ fixation. The basin-wide averages are $14\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$ for the Atlantic, $63\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$ for the Pacific and $26\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$ for the Indian Ocean, giving a global total of $103\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$ (see Supplementary Information). Our own average N₂-fixation rates measured with the bubble-addition method and extrapolated over the Atlantic Ocean (25° N to 45° S), give an input of $14\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$, which is identical to the value for the Atlantic Ocean calculated from the published data. By contrast, our average N₂-fixation rates measured with the dissolution method raise the Atlantic Ocean N₂-fixation rate to $24\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$. If this relative difference applies equally to other ocean basins—where the diazotrophic communities may differ from those of the Atlantic Ocean—the global marine N₂-fixation rate based on direct measurements may have to be increased to $177 \pm 8\,\text{Tg}\,\text{N}\,\text{yr}^{-1}$.

Our study demonstrates that field measurements of N₂ fixation made with the widely applied $^{15}N_2$ bubble-addition method have significantly and variably underestimated N₂-fixation rates in comparison with the dissolution method, with possible implications for the global marine nitrogen budget. Our data reveal regional variations in the magnitude of the underestimate, which may be related to the composition of the diazotrophic community. Specifically, the contribution of diazotrophs other than *Trichodesmium* may have been severely underestimated in previous field studies, leading to a biased view of the key players in this globally important process.

1. Karl, D. *et al.* Dinitrogen fixation in the world's oceans. *Biogeochemistry* **57**, 47–98 (2002).
2. Falkowski, P. G., Barber, R. T. & Smetacek, V. Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**, 200–206 (1998).
3. Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N. & Dunne, J. P. Spatial coupling of nitrogen inputs and losses in the ocean. *Nature* **445**, 163–167 (2007).
4. Altabet, M. A. Constraints on oceanic N balance/imbalance from sedimentary ^{15}N records. *Biogeosciences* **4**, 75–86 (2007).
5. Mahaffey, C., Michaels, A. F. & Capone, D. G. The conundrum of marine N₂ fixation. *Am. J. Sci.* **305**, 546–595 (2005).
6. Codispoti, L. A. An oceanic fixed nitrogen sink exceeding 400 Tg N a⁻¹ vs the concept of homeostasis in the fixed-nitrogen inventory. *Biogeosciences* **4**, 233–253 (2007).
7. Montoya, J. P., Voss, M., Kahler, P. & Capone, D. G. A simple, high-precision, high-sensitivity tracer assay for N₂ fixation. *Appl. Environ. Microbiol.* **62**, 986–993 (1996).
8. Mohr, W., Grosskopf, T., Wallace, D. W. R. & LaRoche, J. Methodological underestimation of oceanic nitrogen fixation rates. *PLoS ONE* **5**, e12583 (2010).
9. Behrenfeld, M. J. & Falkowski, P. G. Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.* **42**, 1–20 (1997).
10. Falkowski, P. G. Evolution of the nitrogen cycle and its influence on the biological sequestration of CO₂ in the ocean. *Nature* **387**, 272–275 (1997).
11. Mills, M. M., Ridame, C., Davey, M., La Roche, J. & Geider, R. J. Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* **429**, 292–294 (2004).
12. Moore, J. K., Doney, S. C., Lindsay, K., Mahowald, N. & Michaels, A. F. Nitrogen fixation amplifies the ocean biogeochemical response to decadal timescale variations in mineral dust deposition. *Tellus B* **58**, 560–572 (2006).
13. Naqvi, S. W. A. *et al.* Increased marine production of N₂O due to intensifying anoxia on the Indian continental shelf. *Nature* **408**, 346–349 (2000).
14. Duce, R. A. *et al.* Impacts of atmospheric anthropogenic nitrogen on the open ocean. *Science* **320**, 893–897 (2008).
15. Gruber, N. & Galloway, J. N. An Earth-system perspective of the global nitrogen cycle. *Nature* **451**, 293–296 (2008).
16. Gruber, N. & Sarmiento, J. L. Global patterns of marine nitrogen fixation and denitrification. *Glob. Biogeochem. Cycles* **11**, 235–266 (1997).
17. Zehr, J. P. *et al.* Unicellular cyanobacteria fix N₂ in the subtropical North Pacific Ocean. *Nature* **412**, 635–638 (2001).
18. Zehr, J. P. *et al.* Globally distributed uncultivated oceanic N₂-fixing cyanobacteria lack oxygenic photosystem II. *Science* **322**, 1110–1112 (2008).

19. Moisander, P. H. *et al.* Unicellular cyanobacterial distributions broaden the oceanic N$_2$ fixation domain. *Science* **327,** 1512–1514 (2010).
20. Langlois, R. J., LaRoche, J. & Raab, P. A. Diazotrophic diversity and distribution in the tropical and subtropical Atlantic Ocean. *Appl. Environ. Microbiol.* **71,** 7910–7919 (2005).
21. Langlois, R. J., Hummer, D. & LaRoche, J. Abundances and distributions of the dominant *nifH* phylotypes in the Northern Atlantic Ocean. *Appl. Environ. Microbiol.* **74,** 1922–1931 (2008).
22. Hamersley, M. R. *et al.* Nitrogen fixation within the water column associated with two hypoxic basins in the Southern California Bight. *Aquat. Microb. Ecol.* **63,** 193–205 (2011).
23. Fernandez, C., Farias, L. & Ulloa, O. Nitrogen fixation in denitrified marine waters. *PLoS ONE* **6,** e20539 (2011).
24. Farnelid, H. *et al.* Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. *PLoS ONE* **6,** e19223 (2011).
25. Halm, H. *et al.* Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. *ISME J.* **6,** 1238–1249 (2012).
26. Luo, Y. W. *et al.* Database of diazotrophs in global ocean: abundances, biomass and nitrogen fixation rates. *Earth Syst. Sci. Data Discuss.* **5,** 47–106 (2012).
27. Moore, C. M. *et al.* Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nature Geosci.* **2,** 867–871 (2009).
28. Subramaniam, A. *et al.* Amazon River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean. *Proc. Natl Acad. Sci. USA* **105,** 10460–10465 (2008).

LETTER

More extreme swings of the South Pacific convergence zone due to greenhouse warming

Wenju Cai[1], Matthieu Lengaigne[2], Simon Borlace[1], Matthew Collins[3,4], Tim Cowan[1], Michael J. McPhaden[5], Axel Timmermann[6], Scott Power[7], Josephine Brown[7], Christophe Menkes[8], Arona Ngari[9], Emmanuel M. Vincent[2] & Matthew J. Widlansky[10]

The South Pacific convergence zone (SPCZ) is the Southern Hemisphere's most expansive and persistent rain band, extending from the equatorial western Pacific Ocean southeastward towards French Polynesia[1,2]. Owing to its strong rainfall gradient, a small displacement in the position of the SPCZ causes drastic changes to hydroclimatic conditions and the frequency of extreme weather events—such as droughts, floods and tropical cyclones—experienced by vulnerable island countries in the region[1–7]. The SPCZ position varies from its climatological mean location with the El Niño/Southern Oscillation (ENSO), moving a few degrees northward during moderate El Niño events and southward during La Niña events[2,5,6]. During strong El Niño events, however, the SPCZ undergoes an extreme swing—by up to ten degrees of latitude toward the Equator—and collapses to a more zonally oriented structure[5] with commensurately severe weather impacts[5,8–11]. Understanding changes in the characteristics of the SPCZ in a changing climate is therefore of broad scientific and socioeconomic interest. Here we present climate modelling evidence for a near doubling in the occurrences of zonal SPCZ events between the periods 1891–1990 and 1991–2090 in response to greenhouse warming, even in the absence of a consensus on how ENSO will change[12–14]. We estimate the increase in zonal SPCZ events from an aggregation of the climate models in the Coupled Model Intercomparison Project phases 3 and 5 (CMIP3[15] and CMIP5) multi-model database that are able to simulate such events. The change is caused by a projected enhanced equatorial warming in the Pacific[16] and may lead to more frequent occurrences of extreme events across the Pacific island nations most affected by zonal SPCZ events.

The SPCZ plays an important part in global circulation and is a major feature of the Southern Hemisphere's climate[1,2]. Its location largely controls rainfall, ocean circulation and tropical cyclogenesis patterns in the South Pacific[5–7]. The western, more equatorial portion of the SPCZ rainfall band is largely controlled by sea surface temperature (SST), whereas its eastern portion is also influenced by extratropical circulation and the subtropical dry zone of the southeastern Pacific[1,17]. As the SPCZ moves northward during El Niño events, countries located within the climatological SPCZ position experience forest fires and droughts[3,4] as well as an increased probability of tropical cyclone hits[5]. In addition, the associated environmental changes affect fisheries[18] and cause coral reef mortality through thermally induced coral bleaching[8–11] across the South Pacific. Observed zonal SPCZ events, characterized by a collapse of the meridional tilt of the rain band, often associated with an equatorward shift of the intertropical convergence zone, have occurred in conjunction with the strongest recorded El Niño events (for example, 1982/83, 1997/98), in which large SST anomalies develop in the central and eastern equatorial Pacific. The effects of these zonal SPCZ events are

much more severe than those from weaker El Niño events, and include massive drought and food shortage[7], unprecedented coral bleaching-induced mortality[9,10] and cyclogenesis in the vicinity of French Polynesia[5], a region not accustomed to such occurrences.

These dramatic effects raise the question as to whether greenhouse warming will change the frequency of zonal SPCZ events. Although many studies have addressed the effects of a projected warming on the Pacific mean state[13,14,19,20], ENSO and its multiplicity[12–14,21,22] and the mean position of the SPCZ[23], the issue of how zonal SPCZ events will change in a warming climate has received little attention. Here we show that greenhouse warming leads to a significant increase in the frequency of such events.

We apply a multivariate signal processing method referred to as empirical orthogonal function (EOF) analysis to deconvolve the spatiotemporal rainfall variability into orthogonal modes, each described by a principal spatial pattern and an associated principal component time series[24]. For observations, we use the satellite-era rainfall data set[25] focusing on austral summer (December to February) when the SPCZ is best developed. The leading pattern (Fig. 1a) features opposite rainfall anomalies between the equatorial and southwestern Pacific around the climatological rain band position. The second pattern (Fig. 1b) is characterized by opposite rainfall anomalies in the equatorial western and central Pacific.

A nonlinear relationship exists between the two associated principal components (time series). In one cluster of events, the principal components are negatively correlated (blue, green and black dots in Fig. 1c) and rainfall anomalies embedded in the two patterns tend to offset each other east of the Date Line, resulting in rainfall variability located in the western Pacific. This cluster encompasses years characterized by the well-known northward and southward shift of the SPCZ during La Niña and moderate El Niño events, respectively (blue and green lines in Fig. 1a). In the second cluster, the principal components are both positive (red dots in Fig. 1c) and display a positive correlation, resulting in large precipitation anomalies east of the Date Line. This cluster consists of the three reported zonal SPCZ events[5] over the 1979–2011 period (1982/83, 1991/1992 and 1997/98, or one in 11 years) where the eastern portion moves equatorward by more than $10°$ in latitude (red line; Fig. 1b). A nonlinear relationship of the second principal component with a historical ENSO index (for example, Niño3.4) is apparent (Supplementary Fig. 1), although we do not distinguish the canonical from the Modoki[22] ENSO. We define a zonal SPCZ event as one for which the first principal component is greater than one standard deviation and the second principal component is greater than zero.

We use this nonlinear behaviour to assess the performance of 17 coupled general circulation models (CGCM) from the CMIP3[15] database forced with historical anthropogenic and natural forcings

[1]CSIRO Marine and Atmospheric Research, Aspendale, Victoria 3195, Australia. [2]Laboratoire d'Océanographie et du Climat: Expérimentation et Approches Numériques (LOCEAN), Tour 45-55 – 4ème étage – Pièce 408, Case 100 – UPMC, 4 Place Jussieu, 75252 Paris Cedex 05, France. [3]College of Engineering Mathematics and Physical Sciences, Harrison Building, Streatham Campus, University of Exeter, Exeter EX4 4QF, UK. [4]Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK. [5]NOAA/Pacific Marine Environmental Laboratory, Seattle, Washington 98115, USA. [6]IPRC, Department of Oceanography, SOEST, University of Hawaii, Honolulu, Hawaii 96822, USA. [7]Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, Victoria 3001, Australia. [8]Institut de Recherche pour le Développement, 101 Promenade Roger Laroque, BP A5 – 98848, Noumea, New Caledonia. [9]Meteorological Service, PO Box 127, Avarua, Rarotonga, Cook Islands. [10]International Pacific Research Center, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA.
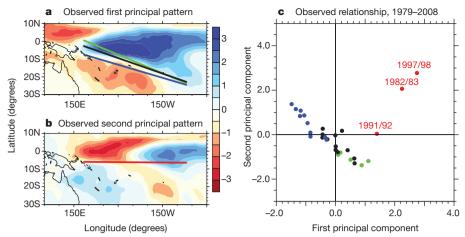
Figure 1 | Principal variability patterns of observed rainfall and their nonlinear relationship. a, b, Spatial patterns obtained by applying a statistical and signal processing method, EOF analysis[24], to a satellite-era rainfall anomaly data from the Global Precipitation Climatology Project version 2[25], focusing on the South Pacific domain (0°–30° S, 160° E–80° W) and the peak season for SPCZ, austral summer (December to February). The associated pattern beyond the domain is obtained by linearly regressing rainfall anomalies onto the time series. The first and second principal spatial patterns account for 47% and 16% of the total variance. The SPCZ position for El Niño (green line), La Niña (blue line) and neutral (black line) states is superimposed in **a**, and the position for zonal SPCZ events (red line) in **b**. The position of the SPCZ is defined as where

the maximum rainfall is greater than 6 mm d^{-1} (ref. 5), and a linear fit is applied. Colour scale at right gives rainfall in mm d^{-1}; blue contours indicate increased rainfall, red contours indicate decreased rainfall per one standard deviation (s.d.) change. **c,** A nonlinear relationship between the associated principal component time series. La Niña, neutral and moderate El Niño years are indicated with blue, black and green dots, respectively. A zonal SPCZ event (red dots) is defined as when the first principal component is greater than one standard deviation, and when the second principal component is greater than zero. The different phases of ENSO events (El Niño and La Niña) are determined from a detrended Niño3.4 index when its amplitude is greater than 0.5 s.d.
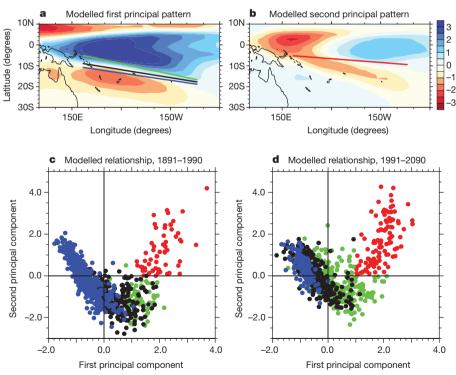
Figure 2 | Multi-model ensemble average of the principal variability patterns of rainfall and their nonlinear relationship from eight CMIP3 CGCMs that are able to produce the nonlinear relationship. a, b, The first (**a**) and second (**b**) principal variability patterns. The multi-model composite SPCZ position over the full 200 years of the control plus climate change periods (see below) for El Niño (green), La Niña (blue) and neutral (black) states is superimposed in **a**, and for zonal SPCZ events (red) in **b**. Colour scale at right gives rainfall in mm d^{-1}; blue contours indicate increased rainfall, and red contours indicate decreased rainfall per one s.d. change. **c, d,** The relationship

between the associated two principal component time series for the control (1891–1990; **c**) and climate change (1991–2090; **d**) periods from the eight CMIP3 CGCMs. La Niña, neutral and moderate El Niño years are indicated with blue, black and green dots, respectively. Variance accounted for by the first and second principal pattern is model-dependent and ranges from 40–53% and 8–15% of the total variance, respectively. An increase in zonal SPCZ events (red dots) from the control to climate change period is evident from a comparison between **c** and **d**. An El Niño (or La Niña) is defined as when the detrended Niño3.4 index has an absolute value greater than 0.5 s.d.

before year 2000, and with a greenhouse gas emission scenario (SRESA2)[15] after year 2000 (see Methods). Nine models fail to generate the nonlinear behaviour (Supplementary Fig. 2). We focus on eight CGCMs that realistically simulate the two principal spatial patterns of rainfall anomalies (Fig. 2a and b), the nonlinear behaviour between the associated principal components (Fig. 2c and d), and the relationship with ENSO (Supplementary Figs 3 and 4). We compare the frequency of zonal SPCZ events in the first (1891–1990) and second (1991–2090) 100-year periods, which we refer to as the 'control' and the 'climate change' periods, respectively, to investigate the influence of greenhouse warming.

Aggregated over these eight CMIP3 CGCMs, the frequency of zonal SPCZ events increases by 81%, from about one event every 14 years in the control period to one every 7.5 years in the climate change period (Fig. 2c and d). This is statistically significant ($P < 0.001$) above the 99% confidence level based on a bootstrap test[26]. The statistical significance is underscored by a strong consensus among CGCMs, with six out of eight CGCMs simulating an increase in occurrences; a sensitivity test to varying definitions of zonal SPCZ events supports the robustness of our result (Supplementary Fig. 3 and Supplementary Table 1).

We conducted an identical analysis on rainfall outputs of 35 available experiments from 20 CMIP5 CGCMs under the historical emissions (before 2005) and a representative concentration pathway (RCP8.5, after 2005) scenario. Aggregated over 15 experiments from eight CMIP5 CGCMs that reproduce the nonlinear behaviour of the SPCZ, there is a similar increase in the occurrences of zonal SPCZ events, from one every 16 years in the control period to one every 7.5 years in the climate change period (Supplementary Table 2 and Supplementary Fig. 5).

We assess the potential impact of well-known biases in simulated climatological SSTs and SPCZ positions[23] using the HadCM3 CGCM (in which biases are corrected through a fixed flux adjustment[13]) in a set of 17 perturbed physics ensemble (PPE) climate change experiments forced with a 1% per year CO_2 increase. In these PPE experiments, perturbations are made to uncertain physical parameters within a single model structure. Aggregated over 12 experiments that produce the nonlinear behaviour of the SPCZ, there is a 214% increase in the frequency of zonal SPCZ events, from one event every 21 years in the control period to one event every 7 years in the climate change period (Supplementary Fig. 6 and Supplementary Table 3). Thus, the conclusions drawn from CMIP3 and CMIP5 CGCMs are not a consequence of the SST biases.
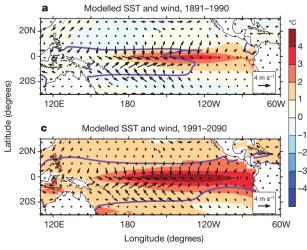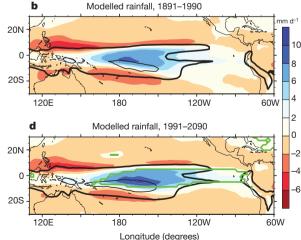
Understanding the relationship between the SPCZ and SST anomalies is crucial for unravelling the mechanism of the increase in zonal SPCZ events in a warming climate. To achieve this, modelled SST, wind and rainfall composite patterns are obtained for zonal SPCZ events (Fig. 3). In both periods, an eastward extension of the Pacific warm pool results in a substantial decrease of the meridional SST gradient south of the Equator, with a rainfall reduction over the western Pacific including northern Australia. In response, eastward-shifted westerly anomalies develop, with a maximum east of the Date Line and south of the Equator, accompanied by strong southeasterlies south of the warm SST anomalies (Fig. 3a and c), as in the observed SST and winds[5,27,28]. Such a wind response increases the moisture convergence east of the Date Line and south of the Equator, displacing the maximum moisture convergence and precipitation axis by up to 10° latitude to the north from its climatological position and contributing to the zonal orientation of the SPCZ (Fig. 3b and d). A zonal SPCZ event therefore occurs when the equatorial western Pacific moves from a state in which the off-equatorial region is warmer than the equatorial region to a condition in which the meridional SST gradient (the off-equatorial minus the equatorial) decreases considerably (Supplementary Figs 7 and 8). An increase in the frequency of zonal SPCZ events may hence arise from more frequent occurrences of a vanishing meridional SST gradient, which may be associated with a change in the occurrences of strong El Niño events. This mechanism also operates for the CMIP5 models for which data are available at the time of writing.

As shown in Supplementary Fig. 9, climate change (greenhouse warming) induces a larger warming rate along the Equator than in the off-equatorial central Pacific, a common feature of the response to global warming[16,29,30]. This pattern results in a slight decrease of the climatological meridional SST gradient in the central Pacific, which translates to an increase in the frequency of austral summers where the meridional SST gradient diminishes (for example, to a value $<0.25\,°C$, Fig. 4a). Such a change occurs in all eight CMIP3 CGCMs (Fig. 4c), and in aggregation, the number of austral summers with a vanishing meridional SST gradient (that is, $<0.25\,°C$) increases by 150%, from one every 15 summers during the control period to one every six summers during the climate change period. In association with this increase, for a given vanishing meridional SST gradient, the number of zonal SPCZ events increases (Fig. 4b).

In addition, there is a 32% increase in strong El Niño events (as measured by a Niño3.4 index greater than a 1.5-s.d. value; the Niño3.4

Figure 3 | Multi-model composites of the circulation fields associated with zonal SPCZ events. Shown are SST (°C), surface winds (m s⁻¹) and rainfall (mm d⁻¹) anomalies referenced to the mean climate of the control period. **a, b**, events over the control period; **c, d**, events over the climate change period. Contours of the composite 28 °C isotherm are superimposed on the SST anomalies (purple), and total rainfall of 6 mm d⁻¹ (thick black contour) and

12 mm d⁻¹ (thin black contour) are superimposed on the rainfall anomalies. The eight CMIP3 CGCM ensemble average shows an increase in rainfall intensity of ∼10% averaged over the SPCZ area during such zonal SPCZ events, from the control period to the climate change period. In areas confined by the green curve, the increase is statistically significant ($P < 0.05$) above the 95% confidence level, based on a t-test comparing the two averages.
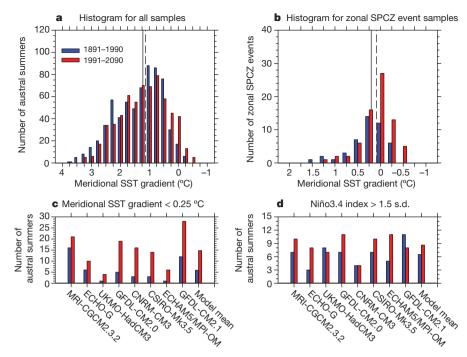
Figure 4 | Multi-model statistics associated with the increase in frequency of zonal SPCZ events. a, Multi-model histogram of the meridional SST gradient in the central equatorial Pacific for the eight CMIP3 CGCMs that are able to produce the nonlinear relationship for the control and climate change period. The meridional SST gradient is defined as the average SST over the off-equatorial region (10° S–5° S, 155° E–120° W) minus the average over the equatorial region (5° S-0°, 155° E–120° W). All 800 years in each period are separated in 0.25 °C bins centred at the tick point for the control (blue) and climate change (red) period. The multi-model median meridional SST gradient

for the control (solid grey line) and the climate change (dashed grey line) periods are indicated. **b**, The same as **a** but for zonal SPCZ events only. Note that the horizontal axes in **a** and **b** are reversed; that is, decreasing go eastward. **c**, Changes in the number of austral summers in which the meridional SST difference virtually vanishes (that is, is <0.25 °C), with all eight CGCMs producing an increase. **d**, Changes in the number of strong El Niño events (the Niño3.4 index greater than a 1.5-s.d. value), with two CGCMs showing a decrease (UKMO-HadCM3 and GFDL-CM2.1).

index is defined as SST anomalies averaged over the region of 170° W–120° W, 5° S–5° N; Supplementary Table 4). This increase, seen in five out of eight CMIP3 CGCMs (Fig. 4d), occurs despite a tendency for a decrease in the total number of El Niño events (defined as when Niño3.4 is greater than a 0.5-s.d. or 0.75-s.d. value), and despite a general lack of consensus on how the El Niño amplitude may change[12,13]. However, only about 50% of the total increase in zonal SPCZ events coincides with a strong El Niño event, suggesting that such a change in El Niño is not necessary for an increase in zonal SPCZ events. Indeed, results from the 12 PPE experiments show a lack of consensus among the PPE experiments, despite a large increase in the frequency of zonal SPCZ events (Supplementary Fig. 10 and Supplementary Table 5), highlighting the fundamental importance of a decrease in the meridional SST gradient.

In summary, increased occurrences of zonal SPCZ events are consistent with the background state change in SST, and support our conclusion of an association with human-induced greenhouse warming, despite a lack of consensus on how ENSO will change. Although the simulated frequency in the control period is comparable to the observed frequency, the latter is based on observations over some three decades and may carry a large uncertainty. Thus, we cannot exclude the possibility that the observed frequency may be higher or lower if an extended observational record was available. With the projected large percentage increase, we expect more frequent occurrences of extreme events such as droughts, floods and tropical cyclones in the Pacific Island nations most affected by zonal SPCZ events.

METHODS SUMMARY

The zonal SPCZ events were diagnosed from observations and CGCMs. We propose an identification method based on the historical events, in which we apply EOF analysis[24] to the satellite-era austral summer rainfall anomalies for the 1979–2008 period over the South Pacific domain (160° E–80° W, 0°–30° S). The rainfall

data[25], Global Precipitation Climatology Project monthly precipitation analysis, is obtained from http://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html. The rainfall anomalies are referenced to the climatological mean of the full period. This yields two leading patterns, one reflecting opposite rainfall anomalies between the equatorial and southwestern Pacific around the climatological rain band position, the other characterized by opposite rainfall anomalies in the equatorial western and central Pacific. The outputs of the analysis are arranged so that the principal component time series have a standard deviation of one, and the differences in variance are expressed in the pattern. Anomalies beyond the South Pacific domain are obtained by a linear regression. A zonal SPCZ event is defined as when the first leading time series is greater than one standard deviation, and when the second leading time series is greater than zero. This definition captures the three observed zonal SPCZ events in the second cluster. We also test the sensitivity of our results to varying definitions (Supplementary Table 1). The method is applied to rainfall anomalies of 17 CMIP3 CGCM simulations, each covering 110 years of a pre-twenty-first-century climate change simulation using historical anthropogenic and natural forcings (1891–2000) and another 90 years (2001–90; the longest common period for the CGCMs) from a future greenhouse warming simulation using the A2 greenhouse gas emission scenario (SRESA2)[15]. The same method is applied to 35 CMIP5 experiments under historical anthropogenic and natural forcings and the RCP8.5 forcing scenario, and 17 PPE experiments forced with historical and a 1% per year CO_2 increase, covering a 200-year period.

1. Kiladis, G. N., Storch, H. V. & Loon, H. V. Origin of the South Pacific Convergence Zone. *J. Clim.* **2**, 1185–1195 (1989).
2. Vincent, D. G. The south Pacific convergence zone (SPCZ): a review. *Mon. Weath. Rev.* **122**, 1949–1970 (1994).
3. Salinger, M. J., Renwick, J. A. & Mullan, A. B. Interdecadal Pacific oscillation and South Pacific climate. *Int. J. Climatol.* **21**, 1705–1721 (2001).
4. Kumar, V. V., Deo, R. C. & Ramachandran, V. Total rain accumulation and rain-rate analysis for small tropical Pacific islands: a case study of Suva, Fiji. *Atmos. Sci. Lett.* **7**, 53–58 (2006).

5. Vincent, E. M. *et al.* Interannual variability of the South Pacific Convergence Zone and implications for tropical cyclone genesis. *Clim. Dyn.* **36,** 1881–1896 (2011).

6. Folland, C. K., Renwick, J. A., Salinger, M. J. & Mullan, A. B. Relative influence of the interdecadal Pacific oscillation and ENSO on the South Pacific Convergence Zone. *Geophys. Res. Lett.* **29,** 1643, http://dx.doi.org/10.1029/2001GL014201 (2002).

7. Hennessy, K. Power, S. & Cambers, G. (eds) *Climate Change in the Pacific: Scientific Assessment and New Research* Vol. 1, *Regional Overview*; Vol. 2, *Country Reports* (Australian Bureau of Meteorology and CSIRO, 2011).

8. Barnett, J. Dangerous climate change in Pacific Islands: food production and food security. *Reg. Environ. Change* **11,** 229–237 (2011).

9. Glynn, P. Widespread coral mortality and the 1982–83 El Niño warming event. *Environ. Conserv.* **11,** 133–146 (1984).

10. Hoegh-Guldberg, O. Climate change, coral beaching and the future of the world's coral reefs. *Mar. Freshwat. Res.* **50,** 839–866 (1999).

11. Mumby, P. J. *et al.* Unprecedented bleaching-induced mortality in *Porites* spp. at Rangiroa Atoll, French Polynesia. *Mar. Biol.* **139,** 183–189 (2001).

12. Guilyardi, E. *et al.* Understanding El Niño in ocean–atmosphere general circulation models: progress and challenges. *Bull. Am. Meteorol. Soc.* **90,** 325–340 (2009).

13. Collins, M. *et al.* The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geosci.* **3,** 391–397 (2010).

14. Ashok, K., Sabin, T. P., Swapna, P. & Murtugudde, R. G. Is a global warming signature emerging in the tropical Pacific? *Geophys. Res. Lett.* **39,** L02701, http://dx.doi.org/10.1029/2011GL050232 (2012).

15. Meehl, G. *et al.* The WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bull. Am. Meteorol. Soc.* **88,** 1383–1394 (2007).

16. Xie, S. P. *et al.* Global warming pattern formation: sea surface temperature and rainfall. *J. Clim.* **23,** 966–986 (2010).

17. Takahashi, K. & Battisti, D. S. Processes controlling the mean tropical Pacific precipitation pattern. Part II: the SPCZ and the southeast Pacific dry zone. *J. Clim.* **20,** 5696–5706 (2007).

18. Lehodey, P., Bertignac, M., Hampton, J., Lewis, A. & Picaut, J. El Niño Southern Oscillation and tuna in the western Pacific. *Nature* **389,** 715–718 (1997).

19. Cane, M. A. *et al.* Twentieth-century sea surface temperature trends. *Science* **275,** 957–960 (1997).

20. Vecchi, G. A. *et al.* Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature* **441,** 73–76 (2006).

21. Yeh, S.-W. *et al.* El Niño in a changing climate. *Nature* **461,** 511–514 (2009).

22. Ashok, K., Behera, S. K., Rao, S. A., Weng, H. & Yamagata, T. El Niño Modoki and its possible teleconnection. *J. Geophys. Res.* **112,** C11007, http://dx.doi.org/10.1029/2006JC003798 (2007).

23. Brown, J. R., Moise, A. F. & Delage, F. P. Changes in the South Pacific Convergence Zone in IPCC AR4 future climate projections. *Clim. Dyn.* **39,** 1–19 (2012).

24. Lorenz, E. N. *Empirical Orthogonal Functions and Statistical Weather Prediction* (Statistical Forecast Project Rep. 1, Department of Meteorology, MIT, 1956).

25. Adler, R. F. *et al.* The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-present). *J. Hydrometeorol.* **4,** 1147–1167 (2003).

26. Austin, P. C. Bootstrap methods for developing predictive models. *Am. Stat.* **58,** 131–137 (2004).

27. Spencer, H. Role of the atmosphere in seasonal phase locking of El Niño. *Geophys. Res. Lett.* **31,** L24104, http://dx.doi.org/10.1029/2004GL021619 (2004).

28. Lengaigne, M., Boulanger, J. P., Menkes, C. & Spencer, H. Influence of the seasonal cycle on the termination of El Niño events in a coupled general circulation model. *J. Clim.* **19,** 1850–1868 (2006).

29. Liu, Z., Vavrus, S., He, F., Wen, N. & Zhong, Y. Rethinking tropical ocean response to global warming: the enhanced equatorial warming. *J. Clim.* **18,** 4684–4700 (2005).

30. Timmermann, A., McGregor, S. & Jin, F.-F. Wind effects on past and future regional sea level trends in the southern Indo-Pacific. *J. Clim.* **23,** 4429–4437 (2010).

Author Contributions W.C. and M.L. conceived the study, directed the analysis and wrote the initial draft of the paper. S.B. performed the analysis. M.C. conducted the perturbed physics ensemble climate change experiments with the HadCM3 model. All authors contributed to interpreting results, discussion of the associated statistical significance, and improvement of the paper.

METHODS

Selection of models. Whereas most of the comprehensive CGCMs are able to simulate the climatological mean SPCZ position and the relationship between the SPCZ and the ENSO[23], their ability to simulate the observed nonlinear relationship between the SPCZ and equatorial warming has not been assessed. An EOF analysis[24] is performed on the full 200 years of austral summer rainfall anomaly data from each of the 17 CMIP3 CGCMs[12]. The name of each CGCM is detailed in Supplementary Table 1. This analysis should yield two leading patterns, one reflecting opposite rainfall anomalies between the equatorial and southwestern Pacific around the climatological rain band position, the other characterized by opposite rainfall anomalies in the equatorial western and central Pacific. The relationship between the EOF1 and EOF2 time series should be nonlinear. A zonal SPCZ event is defined as when the first leading time series is greater than one standard deviation, and when the second leading time series is greater than zero. This definition captures the three observed zonal SPCZ events in the second cluster. Nine CGCMs fail to reproduce the observed feature in terms of a nonlinear relationship between EOF1 and EOF2 time series. These same nine models also poorly simulate the EOF2 pattern with contrasting anomalies over the equatorial western and central Pacific (Supplementary Fig. 2 and Supplementary Table 1). We therefore focus on the remaining eight models that do reproduce the observed nonlinear behaviour. Out of the 35 CMIP5 experiments from 20 CGCMs (available at the time of writing), only eight CGCMs with a total of 15 experiments produce the nonlinear behaviour of the SPCZ (Supplementary Table 2). Likewise, out of 17 PPE experiments, only 12 CGCMs simulate the observed nonlinearity of the SPCZ. Outputs from all experiments run from 1891 to 2090, under historical anthropogenic and natural forcings and then under projected climate scenarios. We derive changes in the occurrence of zonal SPCZ events by comparing the frequency of the first 100 years (control period) from that of the second 100 years (climate change period) in CGCMs that are able to produce zonal SPCZ events.

Statistical significance test. We use a bootstrap method[26] to examine whether the change in frequency of the zonal SPCZ events is statistically significant. The 800-year samples from the eight CMIP3 CGCMs in the control period are re-sampled randomly to construct another 5,000 realizations of 800-year records. During the random re-sampling process, overlapping is allowed, so that any one zonal SPCZ event can be selected again. The standard deviation of the zonal SPCZ frequency in the inter-realization is 7.4 events per 800 years, far smaller than the difference between the control and the climate change periods at 48 events per 800 years. The maximum frequency is 86, far smaller than that in the climate change period, further highlighting the strong statistical significance of the difference between the two periods. Increasing the realizations to 10,000 yields an essentially identical result. This process is repeated for the 15 CMIP5 experiments to create 10,000 realizations of 1,500 samples, and for the 12 PPE experiments to generate 10,000 realizations of 1,200 samples, from their respective control period. In both cases, the standard deviation of the zonal SPCZ frequency in the inter-realization is far smaller that the difference between the climate change and control period, while the maximum frequency of zonal SPCZ events is far smaller than that in the climate change period, indicating a strong statistical significance. The eight CMIP3 CGCM ensemble average shows a ~10% intensity rainfall increase over the SPCZ area, from the control period to the climate change period (Fig. 3d). We use a t-test to compare the two averages and to determine whether a statistically significance above the 95% confidence level is achieved.

LETTER

Reconstructing Native American population history

David Reich[1,2], Nick Patterson[2], Desmond Campbell[3,4], Arti Tandon[1,2], Stéphane Mazieres[3,5], Nicolas Ray[6], Maria V. Parra[3,7], Winston Rojas[3,7], Constanza Duque[3,7], Natalia Mesa[3,7], Luis F. García[7], Omar Triana[7], Silvia Blair[7], Amanda Maestre[7], Juan C. Dib[8], Claudio M. Bravi[3,9], Graciela Bailliet[9], Daniel Corach[10], Tábita Hünemeier[3,11], Maria Cátira Bortolini[11], Francisco M. Salzano[11], María Luiza Petzl-Erler[12], Victor Acuña-Alonzo[13], Carlos Aguilar-Salinas[14], Samuel Canizales-Quinteros[15,16], Teresa Tusié-Luna[15], Laura Riba[15], Maricela Rodríguez-Cruz[17], Mardia Lopez-Alarcón[17], Ramón Coral-Vazquez[18], Thelma Canto-Cetina[19], Irma Silva-Zolezzi[20]†, Juan Carlos Fernandez-Lopez[20], Alejandra V. Contreras[20], Gerardo Jimenez-Sanchez[20]†, Maria José Gómez-Vázquez[21], Julio Molina[22], Ángel Carracedo[23], Antonio Salas[23], Carla Gallo[24], Giovanni Poletti[24], David B. Witonsky[25], Gorka Alkorta-Aranburu[25], Rem I. Sukernik[26], Ludmila Osipova[27], Sardana A. Fedorova[28], René Vasquez[29], Mercedes Villena[29], Claudia Moreau[30], Ramiro Barrantes[31], David Pauls[32], Laurent Excoffier[33,34], Gabriel Bedoya[7], Francisco Rothhammer[35], Jean-Michel Dugoujon[36], Georges Larrouy[36], William Klitz[37], Damian Labuda[30], Judith Kidd[38], Kenneth Kidd[38], Anna Di Rienzo[25], Nelson B. Freimer[39], Alkes L. Price[2,40] & Andrés Ruiz-Linares[3]

The peopling of the Americas has been the subject of extensive genetic, archaeological and linguistic research; however, central questions remain unresolved[1–5]. One contentious issue is whether the settlement occurred by means of a single[6–8] migration or multiple streams of migration from Siberia[9–15]. The pattern of dispersals within the Americas is also poorly understood. To address these questions at a higher resolution than was previously possible, we assembled data from 52 Native American and 17 Siberian groups genotyped at 364,470 single nucleotide polymorphisms. Here we show that Native Americans descend from at least three streams of Asian gene flow. Most descend entirely from a single ancestral population that we call 'First American'. However, speakers of Eskimo–Aleut languages from the Arctic inherit almost half their ancestry from a second stream of Asian gene flow, and the Na-Dene-speaking Chipewyan from Canada inherit roughly one-tenth of their ancestry from a third stream. We show that the initial peopling followed a southward expansion facilitated by the coast, with sequential population splits and little gene flow after divergence, especially in South America. A major exception is in Chibchan speakers on both sides of the Panama isthmus, who have ancestry from both North and South America.

The settlement of the Americas occurred at least 15,000 years ago through Beringia, a land bridge between Asia and America that existed during the ice ages[1–5]. Most analyses of Native American genetic diversity have examined single loci, particularly mitochondrial DNA or the Y chromosome, and some interpretations of these data model the settlement of America as a single migratory wave from Asia[6–8]. We assembled native population samples from Canada to the southern tip of South America, genotyped them on single nucleotide polymorphism (SNP) microarrays, and merged our data with six other data sets. The combined data set consists of 364,470 SNPs genotyped in 52 Native American populations (493 samples; Fig. 1a and Supplementary Table 1), 17 Siberian populations (245 samples; Supplementary Fig. 1 and Supplementary Table 2) and 57 other populations (1,613 samples) (Supplementary Notes).

A complication in studying Native American genetic history is admixture with European and African immigrants since 1492. Cluster analysis[16] shows that many of the samples we examined have some non-native admixture (an average of 8.5%; Fig. 1b and Supplementary Tables 1 and 3). This admixture is a challenge for learning about the historical relationships among the populations, and to address this complication we used three independent approaches. First, we restricted analyses to 163 Native Americans from 34 populations without evidence of admixture (Supplementary Notes). Second, we subtracted the expected contribution of European and African ancestry to the statistics we used to learn about population relationships (Supplementary Notes). Third, we inferred the probability of non-native ancestry at each genomic segment and 'masked' segments with more than a negligible probability of this ancestry (Fig. 1b,
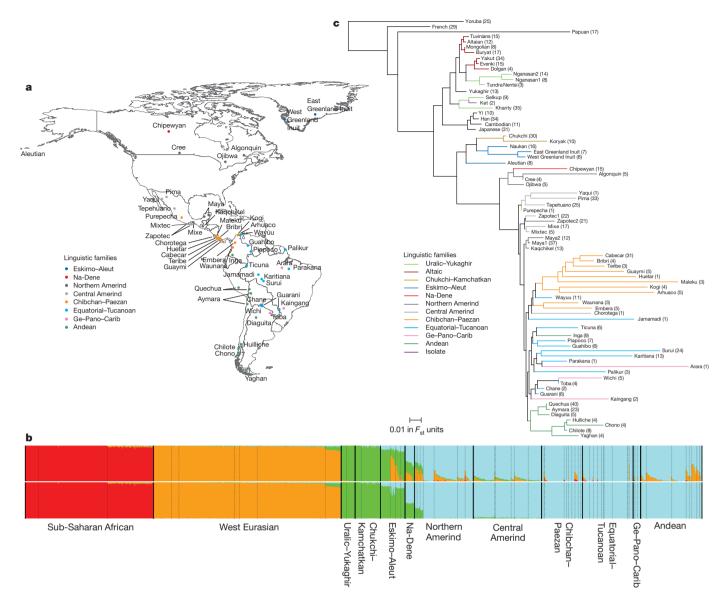
Figure 1 | Geographic, linguistic and genetic overview of 52 Native American populations. a, Sampling locations of the populations, with colours corresponding to linguistic groups. **b,** Cluster-based analysis ($k = 4$) using ADMIXTURE shows evidence of some West-Eurasian-related and sub-Saharan-African-related ancestry in many Native Americans before masking (top), but little afterwards (bottom). Thick vertical lines denote major linguistic groupings, and thin vertical lines separate individual populations.

c, Neighbour-joining tree based on F_{st} distances relating Native American to selected non-American populations (sample sizes in parentheses). Native American and Siberian data were analysed after masking, but consistent trees were obtained on a subset of completely unadmixed samples (Supplementary Fig. 3). Some populations have evidence for substructure, and we represent these as two different groups (for example Maya1 and Maya2).

Supplementary Notes and Supplementary Fig. 2). Our inferences from these three approaches are concordant (Supplementary Figs 3 and 4).

We built a tree (Fig. 1c) using F_{st} distances between pairs of populations, which broadly agrees with geography and linguistic categories[17] (trees based on masked and unmasked data were similar; Supplementary Fig. 3). An early split separates Asians from Native Americans and extreme northeastern Siberians (Chukchi, Naukan, Koryak), which is consistent with studies that have identified pan-American variants shared with some northeastern Siberians[6,7,10,18]. Eskimo–Aleut speakers and far-northeastern Siberians form a cluster that is separated from other Native American populations by a long internal branch. Within America the tree shows a series of splits in an approximate north–south sequence beginning with the Arctic, followed by northern North America, northern/central and southern Mexico and lower Central America/Colombia, and ending in three South American clusters (the Andes, the Chaco region and eastern South America). This pattern of splits is consistent with a north–south population expansion, an

inference that is also supported by the negative correlation between heterozygosity and distance from the Bering Strait ($r = -0.48$, $P = 0.007$). This correlation increases if we use 'least cost distances' that consider the coasts as facilitators of migration[19–21], and persists if we exclude four Native North American populations with ancestry from later streams of Asian gene flow (Supplementary Notes and Supplementary Fig. 5).

Trees provide a simplified model of history that does not accommodate the possibility of gene flow after population separation. Circumstantial evidence that some Native American populations may not fit a simple tree comes from cluster analysis, which infers Siberian-related ancestry in some northern North Americans (Fig. 1b), and from single-locus studies that have identified genetic variants shared between Eurasia and North America that are absent from South America[11,22,23]. The advent of genome-wide data sets has allowed the development of a formal four-population test for whether sets of four populations are consistent with a tree. This test is robust to the

Table 1 | Native Americans descend from at least three streams of Asian gene flow

Population groupings tested	P value for this many Asian streams being enough to explain the data			Minimum number of streams of Asian gene flow needed to explain the data
	1	2	3	
East Greenland Inuit/West Greenland Inuit/First American	$<10^{-9}$	0.64	1	2
East Greenland Inuit/Aleutian/First American	$<10^{-9}$	0.57	1	2
West Greenland Inuit/Aleutian/First American	$<10^{-9}$	0.41	1	2
Chipewyan/East Greenland Inuit/First American	$<10^{-9}$	0.02	1	3
Chipewyan/West Greenland Inuit/First American	$<10^{-9}$	0.006	1	3
Chipewyan/Aleutian/First American	$<10^{-9}$	0.03	1	3
Saqqaq/East Greenland Inuit/First American	$<10^{-9}$	6×10^{-6}	1	3
Saqqaq/West Greenland Inuit/First American	$<10^{-9}$	2×10^{-6}	1	3
Saqqaq/Aleutian/First American	$<10^{-9}$	0.17	1	2
Saqqaq/Chipewyan/First American	$<10^{-9}$	0.29	1	2
Saqqaq/Eskimo–Aleut/Chipewyan/First American	$<10^{-9}$	8×10^{-6}	0.27	3

We use the method described in Supplementary Notes to test formally whether specified groupings of Native American populations are consistent with descending from one, two or three streams of gene flow from Asia. We use 'First American' to refer to a pool of 43 populations from Meso-America southward, and 'Eskimo–Aleut' to refer to a pool of East and West Greenland Inuit and Aleuts. We test either three or four population groupings (when there are three groupings, the maximum number of streams we can reject is two, and so the P value for three streams is always 1). At least two streams of Asian gene flow are required to explain all rows ($P < 10^{-9}$). The Chipewyan, Eskimo–Aleut and First Americans can only be jointly explained by at least three streams. Analysis of the Saqqaq Palaeo-Eskimo (using about sixfold fewer SNPs than for the other analyses) show that the Asian ancestry in this individual has a component that is different from that in First Americans and Greenland Inuit, but indistinguishable from the Chipewyan.

ascertainment bias affecting SNP arrays[24]. For each of the 52 Native American populations in turn, we tested the hypothesis that they conform to the tree: ((test population, southern Native American), (outgroup1, outgroup2)) for 45 pairs of ten Asian outgroups. We used a Hotelling T-test to evaluate whether all four-population test f_4 statistics of this form are consistent with the expectation of zero (Supplementary Notes). The test is not significant for 47 populations, which is consistent with their stemming from the same, presumably first, wave of American settlement; we call this ancestry 'First American' (Table 1). In contrast, four populations from northern North America show highly significant evidence of ancestry from additional streams of gene flow from Asia, subsequent to the initial peopling of America, which we confirm through the Hotelling T-test and a complementary test (Supplementary Notes): East Greenland Inuit ($P < 10^{-9}$), West Greenland Inuit ($P < 10^{-9}$), Aleutian Islanders ($P = 9 \times 10^{-5}$) and Chipewyan ($P < 10^{-9}$). The recently sequenced genome of a 4,000-year-old Saqqaq Palaeo-Eskimo from Greenland[25] also has evidence of ancestry that is distinct from more southern Native Americans ($P = 2 \times 10^{-9}$) (Supplementary Notes).

Examination of the values of the f_4 statistics allows us to infer the minimum number of gene flow events from Asia into America consistent with the data. Each stream of gene flow is expected to produce a distinct vector of f_4 statistics, constituting a 'signature' of how the ancestral migrating population relates to present-day Asian populations. By finding the minimum number of vectors whose linear combinations are necessary to produce the vector observed in each population, we infer that a minimum of three gene flow events from Asia are necessary to explain the data from all Native American populations jointly, including the Saqqaq Palaeo-Eskimo (Supplementary Notes). These three episodes correspond to First American ancestry (distributed throughout the Americas) and to two additional streams of gene flow detected in a subset of northern North Americans (East Greenland Inuit, West Greenland Inuit, Aleutian Islanders, Chipewyan and Saqqaq). Table 1 shows that f_4 statistics in the Inuit and Aleutian islanders are consistent with deriving the non-First-American portions of their ancestry from the same later stream of Asian gene flow, providing support for deep shared ancestry between these linguistically linked groups[12,26]. The Na-Dene-speaking Chipewyan have a different pattern of f_4 statistics from Eskimo–Aleut speakers, implying that they descend at least in part from a separate stream of Asian gene flow ($P < 10^{-9}$ for comparisons with the Greenland Inuit; Table 1). This is consistent with the hypothesis that Na-Dene languages mark a distinct migration from Asia[9,17]. Because we only have data from one Na-Dene-speaking group, an important direction for future work will be to test whether the distinct Asian ancestry that we detect in the Chipewyan is a shared signature throughout Na-Dene speakers. Finally, the Saqqaq[25] have a vector of f_4 statistics consistent with that in the Chipewyan, raising the possibility

that the Saqqaq and Chipewyan both carry genetic material from the same later stream of Asian gene flow into the Americas, postdating the First American migration (Supplementary Notes).

To develop an explicit model for the settlement of the Americas, we used the admixture graph (AG) framework[24]. AGs are generalizations of trees that accommodate the possibility of a limited number of unidirectional gene flow events. They are powerful tools for learning about history because they make predictions about the values of f-statistics (such as f_4) that can be used to test the fit of a proposed model[24] (Supplementary Notes). Figure 2 presents an AG relating selected Native American and Old World populations that is a good fit to the data in the sense that none of the f-statistics predicted by the
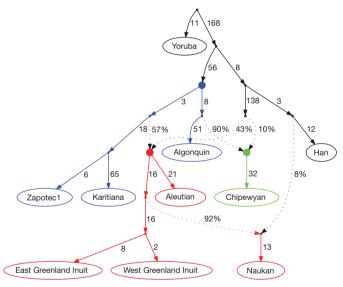
Figure 2 | Distinct streams of gene flow from Asia into America. We present an AG that gives no evidence of being a poor fit to the data and is consistent with three streams of Asian gene flow into America. Solid points indicate inferred ancestral populations, drift on each lineage is given in units proportional to $1,000 \times F_{st}$, and mixture events (dotted lines) are denoted by the percentage of ancestry. The Asian lineage leading to First Americans is the most deeply diverged, whereas the Asian lineages leading to Eskimo–Aleut speakers and the Na-Dene-speaking Chipewyan are more closely related and descend from a common Siberian ancestral population that is a sister group to the Han. The inferred ancestral populations are indicated by filled circles, and the lineages descending from them are coloured: First American (blue), ancestors of the Na-Dene-speaking Chipewyan (green), and Eskimo–Aleut (red). The model also infers a migration of people related to Eskimo–Aleut speakers across the Bering Strait, thus bringing First American genes to Asia (the Naukan are shown, but the Chukchi show a similar pattern; Supplementary Notes).

model are more than three standard errors from what is observed. This supports the hypothesis of three deep lineages in Native Americans: the Asian lineage leading to First Americans is the most deeply diverged, whereas the Asian lineages leading to Eskimo–Aleut speakers and the Na-Dene-speaking Chipewyan are more closely related and descend from a putative Siberian ancestral population more closely related to Han (Fig. 2). We also arrive at the finding that Eskimo–Aleut populations and the Chipewyan derive large proportions of their genomes from First American ancestors: an estimated 57% for Eskimo–Aleut speakers, and 90% in the Chipewyan, probably reflecting major admixture events of the two later streams of Asian migration with the First Americans that they encountered after they arrived (Supplementary Notes). The high proportion of First American ancestry explains why Eskimo–Aleut and Chipewyan populations cluster with First Americans in trees like that in Fig. 1c despite having some of their ancestry from later streams of Asian migration, and explains the observation of some genetic variants that are shared by all Native Americans but are absent elsewhere[6,7,10,18]. We also infer back-migration of populations related to the Eskimo–Aleut from America into far-northeastern Siberia (we obtain an excellent fit to the data when we model the Naukan and coastal Chukchi as mixtures of groups related to the Greenland Inuit and Asians (Fig. 2 and Supplementary Notes)). This explains previous findings of pan-American alleles also in far-northeastern Siberia[6,7,10,18].

We next used AGs to develop a model for the history of populations who derive all their ancestry from the First American migration, with no ancestry from subsequent streams of Asian gene flow. Figure 3 presents an AG we built for 16 selected Native American populations and two outgroups, which is a good fit to the data in that the largest $|Z|$-score for a difference between the observed and predicted f-statistics is 3.2 from among the 11,781 statistics we tested (Supplementary Notes) (The AG of Fig. 3 used masked data; however, a consistent set of relationships is inferred for unadmixed samples (Supplementary Fig. 4).) This model provides a greatly improved statistical fit to the data compared with the tree of Fig. 1c and leads to several novel inferences. First, a relatively large fraction of South American populations fit the AG without a need for admixture events, which we speculate reflects a history of limited gene flow among these populations since their initial divergence. In contrast, only a small fraction of Meso-American populations fit into the AG, which could reflect either a higher rate of migration among neighbouring groups or our denser sampling in Meso-America allowing us to detect more subtle gene flow events. Second, some Meso-American populations have experienced very little genetic drift since divergence from the common ancestral population with South Americans (adding up the genetic drifts along the relevant edges of Fig. 3, we infer $F_{st} = 0.014$ between the Zapotec and a hypothetical population ancestral to all of Central and South America), suggesting that effective population sizes in Meso-America have been relatively large since settlement of the region. Third, the model infers three admixture events consistent with geographic locations and linguistic affiliations (Supplementary Notes). The Inga have both Amazonian and Andean ancestry, which is consistent with their speaking a Quechuan language but living in the eastern Andean slopes of Colombia and thus interacting with groups in the neighbouring Amazonian lowlands. The Guarani stem from two distinct strands of ancestry within eastern South America. The most striking admixture event is in the Costa Rican Cabecar (Fig. 3) and other Chibchan-speaking populations (Supplementary Notes) from the Isthmo-Colombian area. One of the lineages that we detect in these populations occurs definitively within the radiation of South American populations, and so the presence of these populations in lower Central America suggests that there was reverse gene flow across the Panama isthmus after the initial settlement of South America. There has been controversy about whether Chibchan speakers of lower Central America represent direct descendants of the first settlers in the region or more recent migration across the isthmus, and our results support
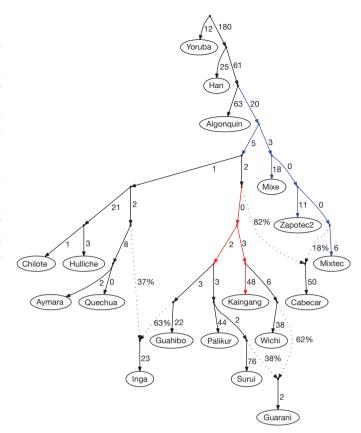

Figure 3 | A model fitting populations of entirely First American ancestry. We show an AG depicting the relationships between 16 selected Native American populations with entirely First American ancestry along with two outgroups (Yoruba and Han). The Colombian Inga are modelled as a mixture of Andean and Amazonian ancestry. The Paraguayan Guarani are fitted as a mixture of separate strands of ancestry from eastern South America. The Central American Cabecar are modelled as a mixture of strands of ancestry related to South Americans and to North Americans, supporting back-migration from South into Central America. The colouring of edges indicates alternative insertion points for the admixing lineages leading to the Cabecar that produce a similar fit to the data in the sense that the χ^2 statistic is within 3.84 of the AG shown. The red colouring shows that the South American lineage contributing to the Cabecar split off after the divergence of the Andean populations, and the blue colouring shows that the other lineage present in the Cabecar diverged before the separation of Andeans. Estimated admixture proportions are shown along the dotted lines, and lineage-specific drift estimates are in units proportional to $1,000 \times F_{st}$.

the view that more recent migration has contributed most of these populations' ancestry[27].

This is the most comprehensive survey of genetic diversity in Native Americans so far. Our analyses show that the great majority of Native American populations—from Canada to the southern tip of Chile—derive their ancestry from a homogeneous 'First American' ancestral population, presumably the one that crossed the Bering Strait more than 15,000 years ago[6–8]. We also document at least two additional streams of Asian gene flow into America, allowing us to reject the view that all present-day Native Americans stem from a single migration wave[6–8], and supporting the more complex scenarios proposed by some other studies[9–15]. In particular, the three distinct Asian lineages we detect—'First American', 'Eskimo–Aleut' and a separate one in the Na-Dene-speaking Chipewyan—are consistent with a three-wave model proposed[9] mostly on the basis of dental morphology and a controversial interpretation of the linguistic data. However, our analyses also document extensive admixture between First Americans and the subsequent streams of Asian migrants, which was not predicted by that model, such that Eskimo–Aleut speakers and the Chipewyan

derive more than half their ancestry from First Americans. Further insights into Native American history will benefit from the application of analyses similar to those performed here to whole-genome sequences and to data from the many admixed populations in the Americas that do not self-identify as native[28–30].

METHODS SUMMARY

The DNA samples we analysed were collected over several decades. For each sample we verified that informed consent was obtained consistent with studies of population history and that institutional approval had been obtained in the country of collection. Ethical oversight and approval for this project was provided by the National Health Service National Research Ethics Service, Central London committee (reference no. 05/Q0505/31). The data set is based on merging Illumina SNP array data newly generated for this study (including 273 Native American samples) with data from six other studies. We applied stringent data curation and validation procedures to the merged data set. We used local ancestry inference software to identify genome segments in each Native American and Siberian sample without evidence of recent European or African admixture, and created a data set that masked segments of potentially non-native origin. Most analyses are performed on the masked data set; however, we confirmed major inferences on a subset of 163 Native American samples that had no evidence of European or African admixture. We used model-based clustering and neighbour-joining trees to obtain an overview of population relationships, and then tested whether proposed sets of four populations were consistent with having a simple tree relationship using the four-population test, which we generalized by means of a Hotelling T-test. We analysed the correlation in allele frequency differences across populations to infer the minimum number of gene flow events that occurred between Asia and America. We fitted the patterns of correlation in allele frequency differences to proposed models of history—AGs—that can incorporate population splits and mixtures.

Full Methods and any associated references are available in the online version of the paper.

1. Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. *The History and Geography of Human Genes* (Princeton Univ. Press, 1994).
2. Meltzer, D. J. *First Peoples in a New World: Colonizing Ice Age America* (Univ. of California Press, 2009).
3. Goebel, T., Waters, M. R. & O'Rourke, D. H. The late Pleistocene dispersal of modern humans in the Americas. *Science* **319,** 1497–1502 (2008).
4. Dillehay, T. D. Probing deeper into first American studies. *Proc. Natl Acad. Sci. USA* **106,** 971–978 (2009).
5. O'Rourke, D. H. & Raff, J. A. The human genetic history of the Americas: the final frontier. *Curr. Biol.* **20,** R202–R207 (2010).
6. Tamm, E. *et al.* Beringian standstill and spread of Native American founders. *PLoS ONE* **2,** e829 (2007).
7. Kitchen, A., Miyamoto, M. M. & Mulligan, C. J. A three-stage colonization model for the peopling of the Americas. *PLoS ONE* **3,** e1596 (2008).
8. Fagundes, N. J. *et al.* Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am. J. Hum. Genet.* **82,** 583–592 (2008).
9. Greenberg, J. H., Turner, C. G. & Zegura, S. L. The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. *Curr. Anthropol.* **27,** 477–497 (1986).
10. Lell, J. T. *et al.* The dual origin and Siberian affinities of Native American Y chromosomes. *Am. J. Hum. Genet.* **70,** 192–206 (2002).
11. Bortolini, M. C. *et al.* Y-chromosome evidence for differing ancient demographic histories in the Americas. *Am. J. Hum. Genet.* **73,** 524–539 (2003).
12. Volodko, N. V. *et al.* Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. *Am. J. Hum. Genet.* **82,** 1084–1100 (2008).
13. Ray, N. *et al.* A statistical evaluation of models for the initial settlement of the American continent emphasizes the importance of gene flow with Asia. *Mol. Biol. Evol.* **27,** 337–345 (2010).
14. de Azevedo, S. *et al.* Evaluating microevolutionary models for the early settlement of the New World: the importance of recurrent gene flow with Asia. *Am. J. Phys. Anthropol.* **146,** 539–552 (2011).
15. Perego, U. A. *et al.* Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr. Biol.* **19,** 1–8 (2009).
16. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19,** 1655–1664 (2009).
17. Ruhlen, M. *A Guide to the World's Languages* (Stanford Univ. Press, 1991).
18. Schroeder, K. B. *et al.* A private allele ubiquitous in the Americas. *Biol. Lett.* **3,** 218–223 (2007).
19. Ray, N. PATHMATRIX: a geographical information system tool to compute effective distances among samples. *Mol. Ecol. Notes* **5,** 177–180 (2005).
20. Wang, S. *et al.* Genetic variation and population structure in native Americans. *PLoS Genet.* **3,** e185 (2007).
21. Yang, N. N. *et al.* Contrasting patterns of nuclear and mtDNA diversity in Native American populations. *Ann. Hum. Genet.* **74,** 525–538 (2010).
22. Brown, M. D. *et al.* mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? *Am. J. Hum. Genet.* **63,** 1852–1861 (1998).
23. Karafet, T. M. *et al.* Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am. J. Hum. Genet.* **64,** 817–831 (1999).
24. Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian population history. *Nature* **461,** 489–494 (2009).
25. Rasmussen, M. *et al.* Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463,** 757–762 (2010).
26. Balter, M. Archaeology. The peopling of the Aleutians. *Science* **335,** 158–161 (2012).
27. Cooke, R. Prehistory of native Americans on the Central American land bridge: Colonization, dispersal, and divergence. *J. Archaeol. Res.* **13,** 129–187 (2005).
28. Wang, S. *et al.* Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet.* **4,** e1000037 (2008).
29. Bryc, K. *et al.* Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc. Natl Acad. Sci. USA* **107** (Suppl 2), 8954–8961 (2010).
30. Wall, J. D. *et al.* Genetic variation in Native Americans, inferred from Latino SNP and resequencing data. *Mol. Biol. Evol.* **28,** 2231–2237 (2011).

Author Contributions D.R., N.B.F., A.L.P. and A.R.-L. conceived the project. D.R., N.P., D.C., A.T., S.M., N.R. and A.R.-L. performed analyses. D.R. and A. R.-L. wrote the paper with input from all the co-authors. A.R.-L. assembled the sample collection, directed experimental work and coordinated the study. All other authors contributed to the collection of samples and data.

Author Information The data analysed here are available for non-profit research on population history under an inter-institutional data access agreement with the Universidad de Antioquia, Colombia; queries regarding data access should be sent to A.R.-L. (a.ruizlin@ucl.ac.uk). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.R. (reich@genetics.med.harvard.edu) or A.R.-L. (a.ruizlin@ucl.ac.uk).

METHODS

DNA samples. The samples analysed here were collected for previous studies over several decades. We reviewed the documentation available for each population to confirm that all samples were collected with informed consent encompassing genetic studies of population history. Institutional approval for use of each set of samples in such research was obtained before this study in the country of collection. Approval for this study was also provided by the National Research Ethics Service, Central London REC 4 (reference no. 05/Q0505/31).

Genotyping. All samples were genotyped by using Illumina arrays, and the data set analysed here is the result of merging data from seven different sources (Supplementary Notes). The genotyping conducted specifically for this study was performed at the Broad Institute of Harvard and Massachusetts Institute of Technology, with the exception of ten Chipewyan samples that were genotyped at McGill University (no systematic differences were observed between these and the five Chipewyan samples genotyped at the Broad Institute). Supplementary Table 3 specifies details for each of the 493 Native American samples. A total of 419 samples were genotyped from genomic DNA, and 74 from whole-genome-amplified material prepared using the Qiagen REPLI-g midi kit.

Data curation. We required more than 95% genotyping completeness for each SNP and sample. We merged the data specifically obtained for this study with six other data sets. We further removed samples that were outliers in principal-component analysis relative to others from their group, showed an excess rate of heterozygotes in comparison with the expected rate from the allele frequencies in the population, or had evidence of being a second-degree relative or closer to another sample in the study (Supplementary Notes). Genetic analyses summarized in the Supplementary Notes found substructure in some populations (Maya, Zapotec and Nganasan); we use labels such as 'Maya1' and 'Maya2' to indicate the subgroups.

Masking of genomic segments containing non-Native American ancestry. For each Native American individual, we used HAPMIX[31] to model their haplotypes with two ancestral panels: first, 'Old World' populations (a pool of 408 Europeans and 130 West Africans) and second, 'Native' populations, a pool of all Native American and Siberian populations. Haplotype phase in the ancestral panel, which is necessary for HAPMIX, was determined by phasing both pools of samples together with Beagle[32]. We masked genome segments that had an expected number of more than 0.01 non-Native American chromosomes according to HAPMIX, thus retaining only segments with an extremely high nominal probability of being homozygous for native ancestry. Multiple analyses reported in Supplementary Information indicate that our masking procedure produces inferences about history that are consistent with those based on unadmixed samples.

Population structure analysis, F_{st} and neighbour-joining tree. We used EIGENSOFT to perform PCA and compute pairwise population F_{st} (ref. 33). Clustering was performed with ADMIXTURE[16]. A neighbour-joining[34] tree based on F_{st} was built with POWERMARKER[35].

Linguistic categories. We used Greenberg's classification[17,36]. We considered using alternative classifications; however, others (for example that in ref. 37) do not propose links between languages at a deep enough level to compare with genetic relationships on a continent-wide scale.

Correlating geography with population diversity. Euclidean distances from the Bering Strait (64.8° N, 177.8° E) and the location of each population (Supplementary Table 1) were calculated by using great arc distances based on a Lambert azimuthal equal-area projection. Least-cost distances between the same points were computed with PATHMATRIX[19], which allowed us to build a spatial cost map incorporating the coastal outline of the Americas. We compared the following coastal/inland relative costs: 1:2, 1:5, 1:10, 1:20, 1:30, 1:40, 1:50, 1:100, 1:200, 1:300, 1:400 and 1:500. We computed a Pearson correlation coefficient between heterozygosity for each population and their least-cost distance from the Bering Strait (Supplementary Notes).

Documentation of at least three streams of gene flow from Asia to America. We used the four-population test to assess whether proposed sets of four populations

were consistent with a tree. For each of 52 test populations, we assessed their consistency with deriving from the same Asian source population as southern Native Americans by studying statistics of the form f_4 (southern Native American, test population; outgroup1, outgroup2), where the two outgroups are the 45 ($=10 \times 9/2$) possible pairs of ten Asian outgroups (Han Chinese and nine Siberian populations with at least ten samples each, and not including the Naukan and Chukchi whom we showed to have some First American ancestry as a result of back-migration across the Bering Strait, making them inappropriate as outgroups (Supplementary Notes)). We applied a Hotelling T-test to assess whether the ensemble of all possible f_4 statistics was consistent with zero after taking into account their correlation structure, resulting in a single hypothesis test for whether the test population was consistent with having the same relationship to the panel of Asian populations as the set of southern Native American samples used as a reference group. We also generalized this test by studying the matrix of all f_4 statistics simultaneously and computing statistics that measured whether the f_4 statistics seen in proposed sets of Native American populations were consistent with deriving from a specified number of Asian migrations. In Supplementary Notes we show that if there have been N distinct streams of gene flow from Asia into the Americas, then the matrix of all possible f_4 statistics can have rank no more than $N - 1$ (ignoring sampling noise). The case $N = 1$ reduces to calculating a Hotelling T^2 statistic. We also developed a likelihood ratio test, generalizing the Hotelling T-test, to evaluate the statistical evidence for larger values of N, allowing us to estimate the minimum number of exchanges between Asia and America that are needed to explain the genetic data.

Admixture graphs. We used the AG framework[24] to fit models of population separation followed by mixture to the data. An AG makes predictions about the correlations in allele frequency differentiation statistics (f-statistics) that will be observed between all pairs, triples and quadruples of populations[24], and these can be compared with the observed values (along with a standard error from a Block Jackknife) to test hypotheses about population relationships (Supplementary Notes). We do not have a formal goodness-of-fit test for whether a given AG fits the data correcting for the number of hypotheses tested and number of degrees of freedom, but use two approximations. First, we examine individual f-statistics, searching for those that are more than three standard errors from expectation indicative of a poor fit. Second, we compute a χ^2 statistic for the match between the observed and predicted f-statistics, taking into account the empirical covariance matrix among the f-statistics computed on the basis of a Block Jackknife. This results in a nominal P value, but it is unclear to us at present whether the empirical covariance matrix that we obtain can be equated with the theoretical covariance matrix that is needed to compute a formal P value. For a fixed graph complexity (number of drift edges and admixture weights), however, we can compare the χ^2 value for different admixture graphs to obtain a formal test for whether some topologies are significantly better fits; this results in the colouring of edges in Fig. 3, which shows alternative insertion points for admixture edges that are equally good fits.

31. Price, A. L. *et al.* Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet.* **5,** e1000519 (2009).
32. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81,** 1084–1097 (2007).
33. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2,** e190 (2006).
34. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425 (1987).
35. Liu, K. & Muse, S. V. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **21,** 2128–2129 (2005).
36. Greenberg, J. H. *Language in the Americas* (Stanford Univ. Press, 1987).
37. Campbell, L. *American Indian languages: the historical linguistics of Native America* (Oxford Univ. Press, 1997).

A biophysical signature of network affiliation and sensory processing in mitral cells

Kamilla Angelo[1,2], Ede A. Rancz[1,3], Diogo Pimentel[1], Christian Hundahl[2,4], Jens Hannibal[4], Alexander Fleischmann[5], Bruno Pichler[3] & Troy W. Margrie[1,3]

One defining characteristic of the mammalian brain is its neuronal diversity[1]. For a given region, substructure, layer or even cell type, variability in neuronal morphology and connectivity persists[2–5]. Although it is well known that such cellular properties vary considerably according to neuronal type, the substantial biophysical diversity of neurons of the same morphological class is typically averaged out and ignored. Here we show that the amplitude of hyperpolarization-evoked sag of membrane potential recorded in olfactory bulb mitral cells is an emergent, homotypic property of local networks and sensory information processing. Simultaneous whole-cell recordings from pairs of cells show that the amount of hyperpolarization-evoked sag potential and current (I_h)[6] is stereotypic for mitral cells belonging to the same glomerular circuit. This is corroborated by a mosaic, glomerulus-based pattern of expression of the HCN2 (hyperpolarization-activated cyclic nucleotide-gated channel 2) subunit of the I_h channel. Furthermore, inter-glomerular differences in both membrane potential sag and HCN2 protein are diminished when sensory input to glomeruli is genetically and globally altered so that only one type of odorant receptor is universally expressed[7]. Population diversity in this intrinsic property therefore reflects differential expression between local mitral cell networks processing distinct odour-related information.

Neurons show a broad array of biophysical properties that profoundly impact the computations that they perform. Even within cell types, diversity in morphology[8], and expression of molecular markers[1] and ion channels[9] is well documented, but whether such variation reflects necessary biological noise[10] or perhaps a functional, dynamic system for regulating excitability at the cellular[11] or even network level remains unclear. The hyperpolarization-activated current (I_h; also observed as a sag potential) is one intrinsic biophysical property that is known to influence the input–output function[12–17] of most principal cell types[18]. In the olfactory bulb, the broad amplitude distribution of I_h-mediated sag potential recorded across the mitral cell population has recently been shown to reflect functional diversity in their input–output responses to *in vivo* stimuli[6].

To explore directly whether cell-to-cell variability in membrane potential sag might reflect differences between functional ensembles of mitral cells, we have taken advantage of the fact that in some brain regions the local architecture facilitates the identification of functionally discrete networks of neurons[19]. This is particularly true for the olfactory bulb, where glomeruli act as information hubs that receive inputs from a unique, homogeneous population of sensory afferents[20] that are integrated by a network of a few hundred interconnected local interneurons and principal mitral and tufted cells. Thus, in a slice preparation, individual mitral cells can be precisely linked to the functional circuit in which they operate *in vivo*[21,22], permitting us to explore whether their intrinsic diversity reflects an emergent property of the functional organization of the olfactory bulb (Fig. 1A).

In wild-type mice, mitral cells can exhibit a hyperpolarization-evoked rebound potential and current indicative of the I_h–mediated sag potential recently described in the rat, both *in vitro* and *in vivo*[6] (Supplementary Fig. 1). The distribution of sag potential amplitude (SPA) recorded across the mitral cell population in mouse is similarly uni-modal ($P < 0.05$) and broad (min = −10.4 mV, max = 30.65 mV, median = 3 mV, mean = 3.43 ± 5.8 mV, $n = 105$ cells, $n = 39$ animals; Fig. 1B and Supplementary Fig. 1). To explore the possibility that such population diversity might reflect differences between local mitral cell networks (Fig. 1A), we performed simultaneous whole-cell recordings of sag from pairs of cells belonging to distinctly different (Fig. 1C, a) or the same glomerular ensemble (Fig. 1C, b). The mean SPA found under these two recording scenarios was not significantly different (inter-glomerular pairs, 3.3 ± 6.7 mV, $n = 52$ cells versus intra-glomerular pairs, 2.45 ± 3.72 mV, $n = 28$ cells ($P = 0.41$); Fig. 1D). For each recorded pair we determined the absolute difference in SPA (Supplementary Fig. 2) and performed a multiple pair-wise comparison, whereby the SPA difference between each cell and all other cells within the same group—excluding its simultaneously recorded 'partner'—was calculated ('pseudo pairs', Fig. 1E). For inter-glomerular pairs of mitral cells, the distribution of SPA difference between recorded and pseudo pairs was similar (recorded: min = 0.03 mV, max = 21.06 mV, median = 3.57 mV, Q1 = 1.87 mV, Q3 = 5.35 mV, $n = 26$ pairs, versus pseudo pairs: min = 0.01 mV, max = 41.05 mV, median = 4.435 mV, Q1 = 1.78 mV, Q3 = 9.46 mV, $n = 1300$ ($P = 0.16$); Fig. 1E, F). This was also the case when comparing inter-glomerular recorded pairs and pseudo pairs extracted from our entire data set ($n = 105$ cells, 5,460 pseudo pairs; Supplementary Fig. 2).

In contrast, the sag potential and the I_h-current amplitude recorded simultaneously from mitral cells belonging to the same glomerular network were virtually indistinguishable (Supplementary Fig. 2). Thus, the difference in the SPA recorded from intra-glomerular pairs was significantly smaller than that determined for intra-glomerular pseudo pairs (recorded SPA difference for intra-glomerular pairs: min = 0 mV, max = 3.59 mV, median = 1.22 mV, Q1 = 0.31, Q3 = 2.1 mV, $n = 14$ pairs, versus intra-glomerular pseudo pairs: min = 0 mV, max = 12.75 mV, median = 4.23 mV, Q1 = 1.99, Q3 = 6.41 mV, $n = 364$ pairs ($P = 0.00002$); Fig. 1E, F and Supplementary Fig. 2), for inter-glomerular recorded pairs ($P = 0.024$; Fig. 1F, G) and pseudo pairs extracted from all cells ($n = 105$ cells; $P = 0.00001$; Fig. 1F and Supplementary Fig. 2). The broad range of sag and I_h-current amplitudes recorded across the bulb therefore reflects differences between individual glomerular circuits, in which it is a homotypic feature of the local mitral cell network (Fig. 1G).

Cell-attached experiments in mitral cell apical dendrites indicate that the I_h current is largest in patches recorded in the very distal region[6]. Thus, the recorded membrane potential sag may reflect activation of HCN channels that are expressed predominantly in the
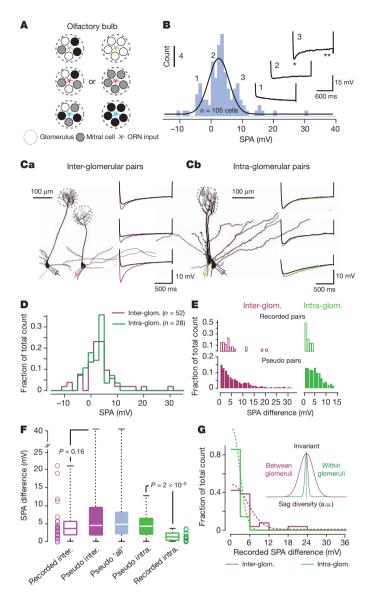
Figure 1 | Diversity of sag potential amplitude within and between mitral cell networks. A, Schematic of two possible scenarios underlying the population diversity of mitral cell sag. Each olfactory glomerulus receives genetically unique input (shown in red, green and blue). Membrane potential sag (shown by different coloured circles) may be hetero- (left) or homogeneously (right) expressed with the glomerulus. ORN, olfactory receptor nerve. **B**, Histogram (fitted with a Gaussian curve) showing the distribution of SPA recorded across the mitral cell population ($n = 105$ cells). Traces, three examples recorded from cells belonging to the indicated bins. The sag potential amplitude is determined as the voltage difference between the peak hyperpolarization (shown by an asterisk) and the steady-state membrane potential (shown by two asterisks). **Ca, b**, Example morphologies of two simultaneously recorded mitral cells projecting to either different (**Ca**) or the same (**Cb**) glomerular networks. The voltage traces show the sag potential recorded in three examples of inter- and intra-glomerular pairs. **D**, Histograms of the SPA for all individual cells belonging to either inter- and intra-glomerular pairs. **E**, Top panel, histograms of recorded SPA differences for inter- ($n = 26$) and intra-glomerular ($n = 14$) pairs. Bottom panel, histograms of SPA differences for inter- and intra-glomerular pseudo pairs. **F**, Box-plot of recorded and pseudo inter-glomerular pairs, pseudo pairs of all recordings, and intra-glomerular pseudo and recorded pairs. Open circles, individual data points of recorded SPA difference. **G**, Histogram of sag amplitude difference for inter- and intra-glomerular recorded pairs fitted with a half Gaussian ($n = 26$ and 14 pairs, respectively; bin size, 3 mV). Inset, Gaussian fits of the recorded data for intra-glomerular and inter-glomerular pairs. a.u., arbitrary units.

dendritic tuft, the site of sensory integration within the glomerulus. To explore this possibility, we performed immunohistochemical staining for the HCN2 subunit that can form both homomeric or heteromeric HCN channels known to mediate the slow I_h (ref. 23) underlying mitral cell sag[6]. Qualitatively, very little HCN2 protein was seen in the granule cell and mitral cell layers[24]. In contrast to the low but homogeneous HCN2 expression in the external plexiform layer[24], we found a high-contrast mosaic staining pattern across the glomerular layer (Fig. 2a). To determine whether HCN2 expression within the glomerulus was postsynaptic to olfactory receptor neuron input we next used a transgenic mouse line that expresses the tau-LacZ transgene in the sensory afferents under the olfactory marker protein (OMP) promoter (OMP-IRES-tau-LacZ mice)[7]. Double-staining experiments against both LacZ and HCN2 showed that the HCN2 protein is predominantly expressed in dendritic compartments within the glomerulus and downstream of the olfactory receptor input (Fig. 2b–d)[25]. Irrespective of the potential contribution of other cell types[24,26], this mosaic pattern of HCN2 expression is consistent with the observation of large I_h currents in the distal apical dendrite and the broad range of sag amplitudes recorded in mitral cells participating in different glomerular networks.

Such glomerular-based SPA and HCN2 expression might reflect network-related homeostatic regulation[27,28] of excitability, in which glomerular differences arise from the processing of functionally and
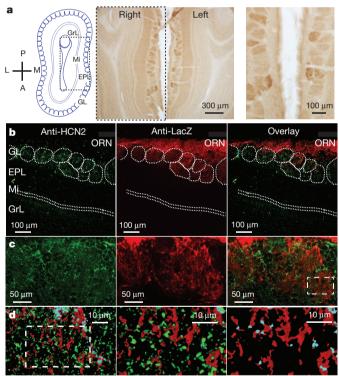
Figure 2 | Glomerular expression of HCN2 in wild-type and OMP-IRES-tau-LacZ mice. a, Left panel, schematic of a horizontal section of an olfactory bulb highlighting its anatomical organization. A, anterior; L, lateral; M, medial; P, posterior. Middle and right panels, HCN2–DAB staining highlighting glomerular diversity. **b**, Left and middle panels, anti-HCN2 and anti-β-galactosidase staining in OMP-IRES-tau-LacZ mice. Right panel, green and red channels shown together for comparison. **c**, High-magnification images of the glomerular layer (colours as in **b**). **d**, Left, higher-magnification images scanned as 30 digital sections, shown as a pseudo three-dimensional image of an approximately 6-µm thick section. Areas of overlap between green and red are shown in light blue and indicate close proximity of the HCN2 and OMP-LacZ signals. Middle and right panels, images of the outlined area shown in the left panel (anti-HCN2 staining (green colour) has been removed for clarity). EPL, external-plexiform layer; GL, glomerular layer; GrL, granule-cell layer; Mi, mitral cell layer; ORN, olfactory receptor nerve layer.

genetically unique subsets of olfactory input[20]. To test this hypothesis we next performed experiments in a transgenic 'monoclonal nose' mouse that expresses the M71 odorant receptor in more than 95% of receptor neurons[7] (M71tg mice, Fig. 3A). We calculated mean pixel intensities of glomeruli in HCN2–DAB-stained M71tg and control animals that revealed significantly different variances ($P = 0.001$) whereby the mean pixel intensity for glomeruli varied less in the M71tg mice than in the control mice ($P = 0.0025$; Fig. 3A, Supplementary Fig. 3). Simultaneous recordings from inter-glomerular pairs of mitral cells in M71tg mice (min = 0.08 mV, max = 6.3 mV, median = 3.3 mV, Q1 = 1.32 mV, Q3 = 3.78 mV; $n = 24$; Fig. 3B, a) also revealed a significantly narrower distribution of SPA difference compared to wild-type and control mice (min = 0.23 mV, max = 21.04 mV, median = 3.4 mV, Q1 = 1.78 mV, Q3 = 10.78 mV, $n = 24$ ($P = 0.004$); Fig. 3B, a and b, C). This reduction in sag amplitude diversity was also notable for pseudo-pair comparisons for all recorded pairs (Fig. 3D, a) and the overall population data set (M71tg mice: min = 0 mV, max = 36.01 mV, median = 3.4 mV, Q1 = 1.6, Q3 = 6.1 mV, $n = 91$ cells, 4,095 comparisons, versus M71 control mice: min = 0 mV, max = 51.4 mV, median = 4.6 mV, Q1 = 2, Q3 = 10.19 mV, $n = 81$ cells, 3,321 comparisons ($P < 4 \times 10^{-47}$); Fig. 3D, a and c). Thus, mitral cells and glomeruli in M71tg mice are more homogeneous in their sag and HCN2 expression profile than those in wild-type and M71 control mice receiving the normal, genetically diverse, array of olfactory receptor neuron input.

Despite the overall reduction in SPA variance in the M71tg mice, the SPA in the intra-glomerular pairs remained more similar (M71tg intra-glomerular pairs: min = 0.2 mV, max = 6.5 mV, median = 0.8 mV, Q1 = 0.3 mV, Q3 = 1.26 mV; $n = 9$, versus M71tg inter-glomerular pairs ($P = 0.04$); Figs 3C and 4a). Indeed, we observed no effect of wholesale expression of the M71 receptor on intra-glomerular sag diversity (M71tg versus control and wild-type mice; $P = 0.61$ and 0.46, respectively; Fig. 4a). Thus, sensory afferent input seems unlikely to be the sole driver of inter-glomerular diversity (Fig. 4b).

Using the hyperpolarization-evoked sag potential as a general proxy, we have identified several organizing principles regarding the population diversity of I_h expression in mitral cells[6,29]. First, this intrinsic property is a biophysical fingerprint of local constellations of mitral cells forming a functionally discrete olfactory network (Fig. 4c). As mitral cells are electrically and exclusively coupled to their intra-glomerular counterparts, co-regulation of the I_h channel and current via gap junctions may contribute to their biophysical similarity[9]. Second, analysis of the M71tg mouse shows directly that this network-affiliation-based signature depends on sensory information processing. The fact that in the M71tg mouse intra-glomerular sag diversity remained more homogeneous than in the overall population also suggests that other factors such as feed-forward and lateral inhibition may contribute to sag regulation at the level of the glomerulus. From a functional perspective, mitral cell F–I curves (frequency of spiking output of neurons (F) in response to varying current injection
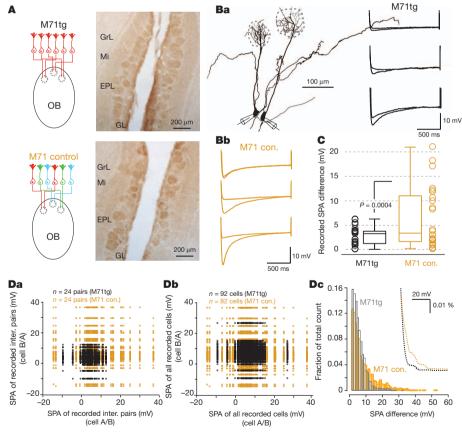
Figure 3 | Glomerular expression of HCN2 and mitral cell sag in M71 monoclonal mice. A, Schematics of the genetic organization of the nose to brain connectivity in M71tg and control mice (left panels). In contrast to the normal genetically diverse ORNs and glomerulus-specific projection pattern in wild-type and control animals, the sensory afferents to all glomeruli in the transgenic mice express the same M71 receptor (red). HCN2–DAB staining (right panels) in horizontal olfactory bulb (OB) slices from an M71tg mouse and a M71 control mouse. **Ba, b**, Example morphologies of two simultaneously recorded mitral cells belonging to distinctly different glomeruli (inter-glomerular pair) in the M71 transgenic mouse. The voltage traces show the sag potential recorded in three different example pairs in M71tg (**Ba**) and control mice (**Bb**). **C**, Box-plot of the SPA difference for inter-glomerular recorded pairs from M71tg and control mice. Open circles, individual data points of recorded SPA difference. **Da, b**, Scatter-plots of SPA for pseudo pairs extracted from paired inter-glomerular recordings in the M71tg and control mice. **Dc**, Histogram of SPA differences for pseudo pairs extracted from all cells recorded in the M71tg ($n = 91$ cells, 4,095 comparisons) and control mice ($n = 81$ cells, 3,321 comparisons; bin size, 1 mV). Inset, histogram spline fits are overlaid and zoomed in to highlight the disparity between the distributions.
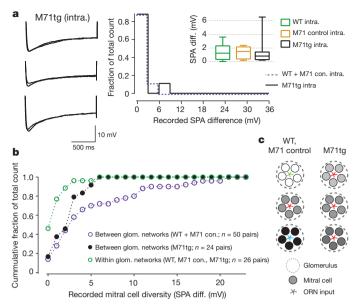
Figure 4 | Population diversity reflects local network membership and sensory processing. a, Left, example membrane-voltage traces showing the sag potential recorded in three different intra-glomerular pairs from the M71tg mouse. Right, histogram of the SPA difference recorded for intra-glomerular pairs in M71tg mice ($n = 9$ pairs) versus M71 control and wild-type (WT) mice (pooled: $n = 17$ pairs). Inset, box-plot of SPA difference for intra-glomerular pairs recorded in wild-type ($n = 14$), M71 control ($n = 3$) and M71tg ($n = 9$) mice. **b**, Summary data plotted as a cumulative histogram of SPA difference of all recorded pairs. **c**, Schematic highlighting the relationship between glomerulus affiliation, sensory input and mitral cell sag diversity.

amplitudes (I)) are known to shift left, from a sigmoid towards a linear operation with increasing sag[6]. We suggest that the glomerular basis of this delineation may therefore reflect a network-based gain control mechanism and contribute to correlated output patterning at the level of mitral cell networks[30]. Irrespective of the cellular mechanisms underlying this glomerular phenomenon, the network-based regulation of this mitral cell property appears fundamental to the organization and function of olfactory bulb circuits.

METHODS SUMMARY

Whole-cell recordings using standard intracellular and extracellular solution were carried out in horizontal olfactory bulb slices (300-μm thick) prepared from wild-type C57Bl/6J and M71 transgenic or control littermate mice aged 4–6 weeks. The I_h current and I_h-mediated sag potential has recently been extensively characterized for mitral cells[6], and the details of the experimental and analytical procedures are provided in the Supplementary Information.

1. Gupta, A., Wang, Y. & Markram, H. Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. *Science* **287,** 273–278 (2000).
2. Brochtrup, A. & Hummel, T. Olfactory map formation in the *Drosophila* brain: genetic specificity and neuronal variability. *Curr. Opin. Neurobiol.* **21,** 85–92 (2011).
3. Reyes, A. *et al.* Target-cell specific facilitation and depression in neocortical networks. *Nature Neurosci.* **1,** 279–285 (1998).
4. Jinno, S. *et al.* Neuronal diversity in GABAergic long-range projections from the hippocampus. *J. Neurosci.* **27,** 8790–8804 (2007).
5. Brown, S. P. & Hestrin, S. Intracortical circuits of pyramidal neurons reflect their long-range axonal targets. *Nature* **457,** 1133–1136 (2009).
6. Angelo, K. & Margrie, T. W. Population diversity and function of hyperpolarization-activated current in olfactory bulb mitral cells. *Scientific Reports* **1,** 1:50 (2011).
7. Fleischmann, A. *et al.* Mice with a ''monoclonal nose'': perturbations in an olfactory map impair odor discrimination. *Neuron* **60,** 1068–1081 (2008).
8. Tsiola, A., Hamzei-Sichani, F., Peterlin, Z. & Yuste, R. Quantitative morphologic classification of layer 5 neurons from mouse primary visual cortex. *J. Comp. Neurol.* **461,** 415–428 (2003).
9. Schulz, D. J., Goaillard, J. M. & Marder, E. Variable channel expression in identified single and electrically coupled neurons in different animals. *Nature Neurosci.* **9,** 356–362 (2006).
10. Ermentrout, G. B., Galan, R. F. & Urban, N. N. Reliability, synchrony and noise. *Trends Neurosci.* **31,** 428–434 (2008).
11. Marder, E. & Goaillard, J. M. Variability, compensation and homeostasis in neuron and network function. *Nature Rev. Neurosci.* **7,** 563–574 (2006).
12. Magee, J. C. Dendritic hyperpolarization-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. *J. Neurosci.* **18,** 7613–7624 (1998).
13. Garden, D. L., Dodson, P. D., O'Donnell, C., White, M. D. & Nolan, M. F. Tuning of synaptic integration in the medial entorhinal cortex to the organization of grid cell firing fields. *Neuron* **60,** 875–889 (2008).
14. George, M. S., Abbott, L. F. & Siegelbaum, S. A. HCN hyperpolarization-activated cation channels inhibit EPSPs by interactions with M-type K$^+$ channels. *Nature Neurosci.* **12,** 577–584 (2009).
15. Nolan, M. F., Dudman, J. T., Dodson, P. D. & Santoro, B. HCN1 channels control resting and active integrative properties of stellate cells from layer II of the entorhinal cortex. *J. Neurosci.* **27,** 12440–12451 (2007).
16. Lüthi, A. & McCormick, D. A. Periodicity of thalamic synchronized oscillations: the role of Ca^{2+}-mediated upregulation of I_h. *Neuron* **20,** 553–563 (1998).
17. Giocomo, L. M., Zilli, E. A., Fransen, E. & Hasselmo, M. E. Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing. *Science* **315,** 1719–1722 (2007).
18. Migliore, M. & Shepherd, G. M. Emerging rules for the distributions of active dendritic conductances. *Nature Rev. Neurosci.* **3,** 362–370 (2002).
19. Woolsey, T. A. & Van der Loos, H. The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. *Brain Res.* **17,** 205–242 (1970).
20. Mombaerts, P. Targeting olfaction. *Curr. Opin. Neurobiol.* **6,** 481–486 (1996).
21. Schoppa, N. E. & Westbrook, G. L. Glomerulus-specific synchronization of mitral cells in the olfactory bulb. *Neuron* **31,** 639–651 (2001).
22. Pimentel, D. O. & Margrie, T. W. Glutamatergic transmission and plasticity between olfactory bulb mitral cells. *J. Physiol.* **586,** 2107–2119 (2008).
23. Robinson, R. B. & Siegelbaum, S. A. Hyperpolarization-activated cation currents: from molecules to physiological function. *Annu. Rev. Physiol.* **65,** 453–480 (2003).
24. Notomi, T. & Shigemoto, R. Immunohistochemical localization of Ih channel subunits, HCN1–4, in the rat brain. *J. Comp. Neurol.* **471,** 241–276 (2004).
25. Santoro, B. *et al.* Molecular and functional heterogeneity of hyperpolarization-activated pacemaker channels in the mouse CNS. *J. Neurosci.* **20,** 5264–5275 (2000).
26. Cadetti, L. & Belluzzi, O. Hyperpolarisation-activated current in glomerular cells of the rat olfactory bulb. *Neuroreport* **12,** 3117–3120 (2001).
27. van Welie, I., van Hooft, J. A. & Wadman, W. J. Homeostatic scaling of neuronal excitability by synaptic modulation of somatic hyperpolarization-activated Ih channels. *Proc. Natl Acad. Sci. USA* **101,** 5123–5128 (2004).
28. Desai, N. S., Rutherford, L. C. & Turrigiano, G. G. Plasticity in the intrinsic excitability of cortical pyramidal neurons. *Nature Neurosci.* **2,** 515–520 (1999).
29. Padmanabhan, K. & Urban, N. N. Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nature Neurosci.* **13,** 1276–1282 (2010).
30. Dhawale, A. K., Hagiwara, A., Bhalla, U. S., Murthy, V. N. & Albeanu, D. F. Non-redundant odor coding by sister mitral cells revealed by light addressable glomeruli in the mouse. *Nature Neurosci.* **13,** 1404–1412 (2010).

Author Contributions K.A. and E.A.R. performed electrophysiological experiments. C.H. and J.H. carried out immunohistochemistry. D.P. performed morphological reconstructions and some initial electrophysiological experiments. B.P. and E.A.R. contributed to data analysis. A.F. generated the transgenic mouse line. K.A. and T.W.M. conceived the project and performed analysis. T.W.M. wrote the paper with input from all other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.W.M. (troy.margrie@nimr.mrc.ac.uk).

LETTER

Activation of specific interneurons improves V1 feature selectivity and visual perception

Seung-Hee Lee[1,2], Alex C. Kwan[1], Siyu Zhang[1], Victoria Phoumthipphavong[1], John G. Flannery[1], Sotiris C. Masmanidis[3], Hiroki Taniguchi[4], Z. Josh Huang[4], Feng Zhang[5], Edward S. Boyden[6], Karl Deisseroth[5] & Yang Dan[1,2]

Inhibitory interneurons are essential components of the neural circuits underlying various brain functions. In the neocortex, a large diversity of GABA (γ-aminobutyric acid) interneurons has been identified on the basis of their morphology, molecular markers, biophysical properties and innervation pattern[1–3]. However, how the activity of each subtype of interneurons contributes to sensory processing remains unclear. Here we show that optogenetic activation of parvalbumin-positive (PV[+]) interneurons in the mouse primary visual cortex (V1) sharpens neuronal feature selectivity and improves perceptual discrimination. Using multichannel recording with silicon probes[4,5] and channelrhodopsin-2 (ChR2)-mediated optical activation[6], we found that increased spiking of PV[+] interneurons markedly sharpened orientation tuning and enhanced direction selectivity of nearby neurons. These effects were caused by the activation of inhibitory neurons rather than a decreased spiking of excitatory neurons, as archaerhodopsin-3 (Arch)-mediated optical silencing[7] of calcium/calmodulin-dependent protein kinase IIα (CAMKIIα)-positive excitatory neurons caused no significant change in V1 stimulus selectivity. Moreover, the improved selectivity specifically required PV[+]

neuron activation, as activating somatostatin or vasointestinal peptide interneurons had no significant effect. Notably, PV[+] neuron activation in awake mice caused a significant improvement in their orientation discrimination, mirroring the sharpened V1 orientation tuning. Together, these results provide the first demonstration that visual coding and perception can be improved by increased spiking of a specific subtype of cortical inhibitory interneurons.

To allow specific activation of PV[+] interneurons, we injected a Cre-inducible adeno-associated virus (AAV) vector containing the ChR2 gene fused in-frame with the coding sequence for enhanced yellow fluorescent protein (eYFP; Supplementary Fig. 1a)[8] into the V1 of a knock-in mouse line expressing Cre recombinase under the PV promoter (PV-ChR2; see Methods)[9,10]. Three weeks after injection, immunostaining confirmed eYFP expression specific to PV[+] neurons (Fig. 1a). To measure the effect of ChR2 activation, we inserted a multichannel silicon probe[4,5] near to the injection site for simultaneous recording from all cortical layers (Supplementary Fig. 1b). After stimulation with blue (473 nm) laser, a small fraction (12 out of 96; 13%) of the neurons showed significant increases in spontaneous firing rate, whereas 43% (41 out of 96) showed significant decreases (Fig. 1b, c and
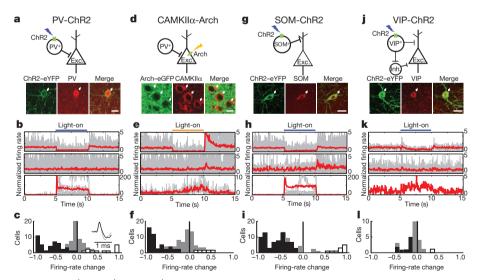
Figure 1 | Optogenetic activation of PV[+], SOM[+] and VIP[+] neurons and silencing of CAMKIIα[+] neurons. a, Fluorescence images of immunostained PV[+] cells (red) expressing ChR2–eYFP (green). Top, schematic illustration of experiment in **a–c**. **b**, Peristimulus time histograms of neurons during 30 repeats of blue-laser stimulation. Top, cells showing a significant firing-rate decrease ($P < 0.01$, bootstrap; $n = 41$); middle, cells without significant change ($n = 43$); bottom, cells with a significant increase ($n = 12$). Grey denotes individual cells and red denotes the average within each group. Blue bar indicates the duration of laser stimulation (5 s). The firing rate of each cell was

normalized to its mean rate over the 5 s before stimulation. **c**, Histogram of firing-rate changes. Black, grey and white bars represent cells showing significant decreases ($P < 0.01$), no change and significant increases, respectively. Inset, spike waveform averaged across cells with significantly decreased (black) or increased (grey) firing. **d–f**, Similar to **a–c** for Arch-mediated silencing of CAMKIIα[+] neurons. **g–i**, Similar to **a–c** for ChR2-mediated activation of SOM[+] neurons. **j–l**, Similar to **a–c** for ChR2-mediated activation of VIP[+] neurons. Exc., excitatory neuron; Inh., inhibitory neuron. Thunderbolt marks denote laser stimulation. Scale bars, 15 μm.

[1]Division of Neurobiology, Department of Molecular and Cell Biology, Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720, USA. [2]Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA. [3]Broad Fellows Program in Brain Circuitry, Kavli Nanoscience Institute, California Institute of Technology, Pasadena, California 91125, USA. [4]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. [5]Department of Bioengineering, Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA. [6]Media Laboratory, Biological Engineering, Brain and Cognitive Sciences, and McGovern Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Supplementary Fig. 1c). The cells with increased firing showed narrower spike waveforms than those with decreased firing (Fig. 1c, inset; peak–trough width, 0.31 ± 0.12 versus 0.44 ± 0.09 ms; mean $\pm$ s.d.), indicating that the PV^+ interneurons activated directly by laser were mostly fast-spiking cells[11]. The decreased spiking of other neurons is probably caused by increased inhibition from the activated PV^+ neurons.

We then measured orientation tuning and direction selectivity of the neurons using drifting grating stimuli, both with and without ChR2 activation. We found that PV^+ neuron activation caused considerable changes in the tuning of other neurons. In addition to the overall firing-rate decrease, we also observed a marked reduction in orientation tuning width and an increase in direction selectivity (Fig. 2a). To quantify these effects, we fitted each tuning curve with a double-Gaussian function (Supplementary Fig. 2a). For all the well-tuned neurons (the tuning curves of which with and without ChR2 activation were well fitted; see Methods), ChR2 activation caused a significant decrease in tuning bandwidth (σ; $P < 10^{-4}$, paired t-test; Fig. 2b) and an increase in direction selectivity index (DSI; $P = 0.007$; Fig. 2c). However, the preferred orientation (θ_0) remained relatively constant (Fig. 2d). Interestingly, all of the putative PV^+ neurons (Fig. 1c, white bars) showed poor tuning even before laser stimulation, consistent with previous findings[12,13]. Although most of the poorly fitted tuning curves

showed low signal-to-noise ratio (Supplementary Fig. 2b, c, examples 1, 2), a few had relatively reliable tuning curves showing several irregular peaks (Supplementary Fig. 2b, c, examples 3–6). Laser stimulation also sharpened the tuning of some of these cells, although the degree of sharpening was not well quantified by curve fitting.

The decrease in σ was correlated with the reduction in firing rate (Supplementary Fig. 3e), suggesting that the enhanced selectivity depends on the increase in inhibition. However, we also wondered whether an overall reduction in cortical firing is sufficient to enhance the selectivity without interneuron activation. As an alternative method to decrease cortical spiking, we expressed the light-activated proton pump Arch[7] in $CAMKII\alpha^+$ excitatory neurons by injecting AAV (Supplementary Fig. 1a) into the V1 of the CAMKIIα-Cre mice[14] (CAMKIIα-Arch; Fig. 1d). Arch-mediated partial silencing of $CAMKII\alpha^+$ neurons caused decreases in the spontaneous (Fig. 1e, f and Supplementary Fig. 1c) and visually evoked (Supplementary Fig. 1c, d) firing rates, similar to PV^+ neuron activation. However, $CAMKII\alpha^+$ neuron silencing caused no significant change in σ ($P = 0.75$; Fig. 2f) or DSI ($P = 0.68$; Fig. 2g). In addition, there was no significant correlation between the firing-rate reduction and change in σ or DSI (Supplementary Fig. 3f, j).

Next, we wanted to test whether the enhanced selectivity required a general increase in inhibition or specific activation of PV^+ interneurons.
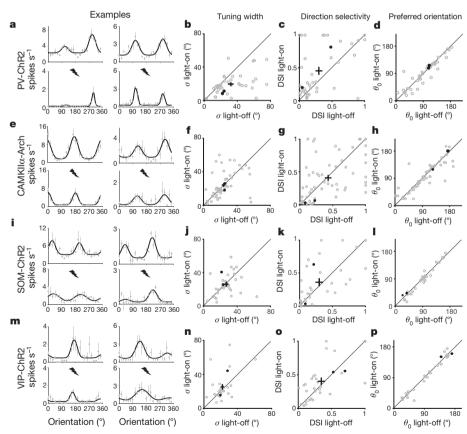
Figure 2 | PV^+ activation enhances V1 stimulus selectivity. a, Tuning curves of two cells, each box for one cell. Grey dots denote measured firing rates and black line denotes fitted curve. Top plot depicts no laser; bottom plot depicts laser stimulation. **b–d,** Population summary of ChR2-mediated changes in stimulus selectivity ($n = 41$). Each circle represents one cell and cross represents population average. **b,** Tuning width (σ), light-off, $32.1 \pm 2.9°$ and light-on, $19.7 \pm 1.9°$; 22% individual neurons showed a significant decrease at $P < 0.01$ (bootstrap) and none showed a significant increase. **c,** DSI, light-off, 0.30 ± 0.05 and light-on, 0.45 ± 0.06; 11% of neurons showed a significant increase and none showed a significant decrease. **d,** Preferred orientation (θ_0), median difference between light-on and light-off, $6.2°$. For neurons in **b–d,** the laser reduced visually evoked firing rate from 5.5 ± 0.6 to 2.8 ± 0.5 spikes s^{-1}

($P < 10^{-4}$, paired t-test). Filled circles, example cells in **a. e–h,** Similar to **a–d** for $CAMKII\alpha^+$ silencing. Mean firing rate reduced from 3.0 ± 0.4 to 2.2 ± 0.4 spikes s^{-1} ($n = 56$; $P = 0.003$). **f,** σ, light-off, $23.6 \pm 1.8°$ and light-on, $24.3 \pm 2.0°$. No individual cell showed a significant change. **g,** DSI, light-off, 0.43 ± 0.04 and light-on, 0.41 ± 0.04. **h,** θ_0, median difference, $5.4°$. **i–l,** Similar to **a–d** for SOM-ChR2 mice. Firing rate reduced from 3.4 ± 0.6 to 1.9 ± 0.4 spikes s^{-1} ($P < 10^{-3}$). **j,** σ, light-off, $26.9 \pm 2.1°$ and light-on, $26.2 \pm 2.3°$, $n = 33$. No individual cell showed a significant change. **k,** DSI, light-off, 0.29 ± 0.05 and light-on, 0.36 ± 0.06. **l,** θ_0, median difference, $5.7°$. **m–p,** Similar to **a–d** for VIP-ChR2 mice ($n = 31$). **n,** σ, light-off, $21.8 \pm 1.8°$ and light-on, $25.5 \pm 2.8°$. **o,** DSI, light-off, 0.35 ± 0.05 and light-on, 0.41 ± 0.05. **p,** θ_0, median difference, $5.6°$. All data are mean $\pm$ s.e.m.

In addition to PV$^+$ neurons, somatostatin (SOM)$^+$ and vasointestinal peptide (VIP)$^+$ neurons constitute two other major subtypes of GABAergic interneurons[2]. To test their roles in shaping V1 selectivity, we induced cell-type-specific expression of ChR2–eYFP in SOM-Cre or VIP-Cre mice (Fig. 1g, j). In the SOM-ChR2 mice, the laser also caused increased spiking of a few neurons (11 out of 91) but decreased firing in most of the neurons (71 out of 91) (Fig. 1h, i, and Supplementary Fig. 1c, d). In fact, compared with the PV-ChR2 group, a higher percentage of neurons showed decreased firing. However, SOM$^+$ activation caused no significant effect on either σ ($P = 0.79$; Fig. 2i, j) or DSI ($P = 0.17$; Fig. 2i, k). In VIP-ChR2 mice, laser stimulation also induced no significant change in σ ($P = 0.16$; Fig. 2m, n) or DSI ($P = 0.16$; Fig. 2m, o), and only a mild change in firing rate (Fig. 1k, l and Supplementary Fig. 1c, d). The change in σ showed a negative correlation with firing-rate change (Supplementary Fig. 3h), opposite to PV$^+$ activation (Supplementary Fig. 3e). This may be because the VIP$^+$ neurons also innervate inhibitory interneurons[15], thus causing both inhibition and disinhibition. In addition to σ, tuning measured by orientation selectivity index was also improved by PV$^+$ activation, but not by CAMKIIα^+ silencing or SOM$^+$ or VIP$^+$ activation (Supplementary Fig. 4).

Even among the PV-ChR2, SOM-ChR2 and CAMKIIα-Arch groups, the degrees of firing-rate suppression were not identical. To ensure that the difference in tuning width reduction was not caused by differences in firing-rate suppression, we selected neurons from each group that showed intermediate levels of suppression (firing-rate change, -0.75 to -0.3; Supplementary Fig. 5a). Although the spike-rate suppression is well matched across groups within this range (mean rate change, -0.51 (PV), -0.50 (SOM), -0.50 (CAMKIIα)), laser stimulation caused a significant decrease in σ in the PV (-0.19 ± 0.05; mean $\pm$ s.e.m.) but not the SOM (0.01 ± 0.09) or CAMKIIα (-0.05 ± 0.03) groups. The magnitudes of reduction in σ were significantly different between the PV and SOM or CAMKIIα groups (Supplementary Fig. 5b). We also noticed that neurons in these groups showed different initial tuning widths ($\sigma_{\text{light-off}}$). To ensure that the difference in tuning width reduction was not caused by the difference in $\sigma_{\text{light-off}}$, we selected neurons with $\sigma_{\text{light-off}}$ falling within two ranges: the broad (25–$50°$) and narrow (0–$25°$) ranges. Within each range, the median values of $\sigma_{\text{light-off}}$ were not significantly different across groups, but laser stimulation caused a significant decrease in σ in the PV-ChR2 group (broad, -0.28 ± 0.09; narrow, -0.11 ± 0.05) and not in any other group (Supplementary Fig. 5c, d). Furthermore, within each range the amount of decrease in σ was significantly different between the PV group and other groups.

Although the lack of effect of VIP$^+$ activation on σ is not surprising given its moderate effect on firing rate, what underlies the difference between SOM$^+$ and PV$^+$ neurons is not obvious. Tuning of a sensory neuron depends on its firing rate versus input current (F–I) curve: subtractive modulation of the F–I curve causes sharpening of tuning, whereas divisive modulation has no effect[16]. We therefore tested the effects of SOM$^+$ and PV$^+$ activation on neuronal F–I functions using in vivo whole-cell recording (see Methods). For each neuron we measured firing rate versus injected depolarizing current. We found that PV$^+$ activation caused a large increase in spiking threshold ($P = 6 \times 10^{-4}$, paired t-test) without significantly changing the slope ($P = 0.34$; Fig. 3a–d), but SOM$^+$ activation caused a marked reduction of the slope ($P = 0.02$) without a significant change in threshold ($P = 0.18$; Fig. 3e–h). These effects may be related to the different subcellular targeting of PV$^+$ and SOM$^+$ neurons onto pyramidal neurons (perisomatic versus dendritic)[17]. When we plotted the median response of each group at each orientation with versus without laser stimulation, we also found a strong subtractive effect of PV$^+$ activation but a primarily divisive effect of SOM$^+$ activation (Supplementary Fig. 6).

Changes in V1 orientation tuning can, in principle, affect perceptual discrimination[18,19]. To test the functional consequence of sharpened tuning, we applied laser stimulation to awake PV-ChR2 mice while they performed orientation discrimination. The head-restrained mice were trained on a go/no-go task and rewarded for licking in response to a grating at one of two orientations[20] (see Methods, Fig. 4a, b and Supplementary Movie 1). For an easy discrimination task (difference between go and no-go orientations $\Delta\theta = 90°$), the performance measured by the divergence between hit and false-alarm rates (Fig. 4c) and discriminability (d'; Fig. 4d) improved systematically over days. As expected, d' increased monotonically with $\Delta\theta$ (Fig. 4e). After the performance reached a steady state, we tested the effect of laser and found a significant improvement in d' at a range of $\Delta\theta$ (Fig. 4f and Supplementary Fig. 7a–c). To ensure that the improvement was caused by optical activation of PV$^+$ neurons rather than through the retinal photoreceptors activated by scattered laser, we performed a control experiment in mice that were not virus-infected and found no significant increase in d' (Fig. 4g). Separate electrophysiological experiments in awake mice showed that PV$^+$ activation also caused a significant sharpening of tuning ($P < 0.05$; Supplementary Fig. 8), similar to that in anaesthetized mice (Fig. 2a, b). Thus, the sharpening of V1 orientation tuning induced by PV$^+$ activation can facilitate orientation discrimination. Unlike the learning curve shown in Fig. 4d, in which d' improved steadily over 5–10 days of training, the effect of PV$^+$ activation showed no systematic change over several
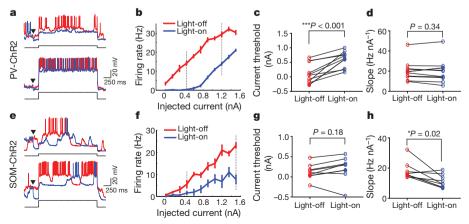
Figure 3 | Effects of PV$^+$ and SOM$^+$ activation on F–I function. a, Example traces showing neuronal spiking evoked by current injection. Red indicates voltage trace without laser and blue indicates voltage trace with laser. Arrowhead indicates laser onset (laser offset, 800 ms after current injection; not shown). Current amplitude, 0.45 nA (upper) and 1.2 nA (lower). **b,** F–I curve of the cell shown in **a.** Red denotes firing rate without laser and blue denotes firing rate with laser. Dashed lines indicate current amplitudes shown in **a. c, d,** Summary of threshold (lowest current that evokes spiking) and F–I slope with and without laser. Each symbol represents one cell ($n = 10$). **e–h,** Similar to **a–d** for SOM$^+$ activation ($n = 8$). All data are ± s.e.m.
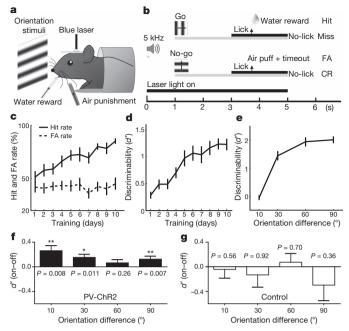
Figure 4 | PV$^+$ activation improves perceptual discrimination. a, Schematic of behavioural experiment. **b,** Task design. Grey bar indicates the duration of visual stimulation and black bar indicates the response window. CR, correct rejection; FA, false alarm. **c,** Changes in hit and FA rates of PV-ChR2 mice over the training period ($n = 25$ mice). **d,** Changes in d' over training period. **e,** d' versus $\Delta\theta$. Data in **c–e** were collected without laser stimulation. **f,** Laser-induced changes in d' in PV-ChR2 mice, significant at $\Delta\theta = 10°$ ($P = 0.008$, Wilcoxon signed-rank test, 0.032 after Bonferroni correction; $n = 9$ mice), $\Delta\theta = 30°$ ($P = 0.011$; $n = 19$) and $\Delta\theta = 90°$ ($P = 0.007$; $n = 25$). **g,** Similar to **f** but for control mice ($n = 10$). All data are mean ± s.e.m.

days of testing (Supplementary Fig. 7d), suggesting that the effect was not caused by learning of a distorted perception owing to PV$^+$ activation.

Previous pharmacological experiments have shown that the endogenous level of GABAergic inhibition is necessary for orientation tuning and direction selectivity[21,22]. Our study showed that increasing inhibitory neuron activity above the normal level can further sharpen cortical feature selectivity and improve perceptual discrimination. A recent study showed that PV$^+$ activation only moderately affects the tuning of V1 neurons[23]. This is probably caused by the relatively low level of PV$^+$ activation compared with our study, as the increase in stimulus selectivity is strongly correlated with the degree of firing-rate suppression (Supplementary Fig. 3e, i). Interestingly, studies in awake, behaving primates have shown that increases in task difficulty[24] and attention[25] are accompanied by a preferential increase in the firing of fast-spiking neurons, many of which are likely to be PV$^+$ interneurons[11]. This suggests that the level of inhibitory activity can be dynamically regulated by top-down mechanisms to meet changing behavioural demands.

A long-standing debate on the mechanism for orientation tuning is whether the excitatory and inhibitory inputs show similar tuning[26]. In the rodent V1, whether inhibitory interneurons are well tuned has also been controversial[12,13,27]. Our results show that an increase in the untuned activity of inhibitory interneurons is sufficient to sharpen orientation tuning, consistent with previous theoretical prediction[28]. Furthermore, although driving either PV$^+$ or SOM$^+$ interneurons markedly suppressed cortical spiking, only PV$^+$ activation could sharpen orientation tuning. This specificity may be attributable to the powerful perisomatic inhibition exerted by PV$^+$ neurons[1,29]. Recent studies have demonstrated the importance of PV$^+$ neurons in gating developmental plasticity[30] and generating gamma oscillations[9,10]. Our results show that these interneurons also have a unique role in visual coding and perception.

METHODS SUMMARY

AAV was injected into the V1 of adult (P40-60) PV-Cre, CAMKIIα-Cre, SOM-Cre or VIP-Cre mice. For recording and behavioural training in awake mice, the head plate was implanted in the same surgery as the virus injection; recording or training was performed 2–6 weeks after surgery. For ChR2 (Arch) activation, an optic fibre coupled to blue (yellow/green) laser was placed on top of the injection site. For measuring orientation tuning, eight repeats of drifting sinusoidal grating (spatial frequency, 0.04 cycles per degree; temporal frequency, 2 Hz; 100% contrast) were presented in 24 directions (0–360°) in pseudorandom sequence. Blocks of trials with or without laser were interleaved. To quantify tuning, we fitted each tuning curve by double-Gaussian function. To test orientation discrimination, mice were trained to discriminate between go and no-go stimuli (drifting gratings at different orientations) for a water reward[20].

Full Methods and any associated references are available in the online version of the paper.

1. Markram, H. *et al.* Interneurons of the neocortical inhibitory system. *Nature Rev. Neurosci.* **5,** 793–807 (2004).
2. Xu, X. M., Roby, K. D. & Callaway, E. M. Immunochemical characterization of inhibitory mouse cortical neurons: three chemically distinct classes of inhibitory cells. *J. Comp. Neurol.* **518,** 389–404 (2010).
3. Ascoli, G. A. *et al.* Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nature Rev. Neurosci.* **9,** 557–568 (2008).
4. Blanche, T. J., Spacek, M. A., Hetke, J. F. & Swindale, N. V. Polytrodes: high-density silicon electrode arrays for large-scale multiunit recording. *J. Neurophysiol.* **93,** 2987–3000 (2005).
5. Du, J. G. *et al.* High-resolution three-dimensional extracellular recording of neuronal activity with microfabricated electrode arrays. *J. Neurophysiol.* **101,** 1671–1678 (2009).
6. Deisseroth, K. Optogenetics. *Nature Methods* **8,** 26–29 (2011).
7. Chow, B. Y. *et al.* High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* **463,** 98–102 (2010).
8. Zhang, F. *et al.* Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. *Nature Protocols* **5,** 439–456 (2010).
9. Sohal, V. S., Zhang, F., Yizhar, O. & Deisseroth, K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459,** 698–702 (2009).
10. Cardin, J. A. *et al.* Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* **459,** 663–667 (2009).
11. Toledo-Rodriguez, M. *et al.* Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. Cortex* **14,** 1310–1327 (2004).
12. Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* **67,** 858–871 (2010).
13. Ma, W. P. *et al.* Visual representations by cortical somatostatin inhibitory neurons–selective but with weak and delayed responses. *J. Neurosci.* **30,** 14371–14379 (2010).
14. Tsien, J. Z. *et al.* Subregion- and cell type-restricted gene knockout in mouse brain. *Cell* **87,** 1317–1326 (1996).
15. Dávid, C., Schleicher, A., Zuschratter, W. & Staiger, J. F. The innervation of parvalbumin-containing interneurons by VIP-immunopositive interneurons in the primary somatosensory cortex of the adult rat. *Eur. J. Neurosci.* **25,** 2329–2340 (2007).
16. Ayaz, A. & Chance, F. S. Gain modulation of neuronal responses by subtractive and divisive mechanisms of inhibition. *J. Neurophysiol.* **101,** 958–968 (2009).
17. Mehaffey, W. H., Doiron, B., Maler, L. & Turner, R. W. Deterministic multiplicative gain control with active dendrites. *J. Neurosci.* **25,** 9968–9977 (2005).
18. Kang, K. J., Shapley, R. M. & Sompolinsky, H. Information tuning of populations of neurons in primary visual cortex. *J. Neurosci.* **24,** 3726–3735 (2004).
19. Schoups, A., Vogels, R., Qian, N. & Orban, G. Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412,** 549–553 (2001).
20. Andermann, M. L. K. e. r. l. i. n. A. M. &. R. e. i. d. R. C. Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing. *Front. Cell. Neurosci.* **4,** 3 (2010).
21. Sillito, A. M. Effectiveness of bicuculline as an antagonist of GABA and visually evoked inhibition in the cat's striate cortex. *J. Physiol. (Lond.)* **250,** 287–304 (1975).
22. Nelson, S., Toth, L., Sheth, B. & Sur, M. Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science* **265,** 774–777 (1994).
23. Atallah, B. V., Bruns, W., Carandini, M. & Scanziani, M. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73,** 159–170 (2012).
24. Chen, Y. *et al.* Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nature Neurosci.* **11,** 974–982 (2008).
25. Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* **55,** 131–141 (2007).
26. Ferster, D. & Miller, K. D. Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* **23,** 441–471 (2000).

27. Runyan, C. A. *et al.* Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron* **67,** 847–857 (2010).
28. Troyer, T. W., Krukowski, A. E., Priebe, N. J. & Miller, K. D. Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. *J. Neurosci.* **18,** 5908–5927 (1998).
29. Di Cristo, G. *et al.* Subcellular domain-restricted GABAergic innervation in primary visual cortex in the absence of sensory and thalamic inputs. *Nature Neurosci.* **7,** 1184–1186 (2004).
30. Hensch, T. K. Critical period plasticity in local cortical circuits. *Nature Rev. Neurosci.* **6,** 877–888 (2005).

METHODS

AAV vectors. ChR2 fused to eYFP (ChR2–eYFP) and Arch fused to enhanced green fluorescent protein (eGFP) (Arch–eGFP) were cloned into pAAV-MCS (Stratagene) in an antisense direction flanked by a pair of canonical *loxP* sites and a pair of mutated *lox2272* sites. AAV particles (serotype 2) were produced by co-transfection of packaging plasmids into HEK293T cells, and cell lysates were fractionated by iodixanol gradient ultracentrifugation[31]. Viral particles were further purified from the crude fraction by heparin affinity column (HiTrap Heparin HP Columns; GE Healthcare), desalted and concentrated with Amicon Ultra Centrifugal Filter (100K, Millipore) (1×10^{12}–1×10^{14} particles ml^{-1} in PBS buffer). AAV serotype 2/2 was used for ChR2 and serotypes 2/2 and 2/8 were used for Arch. All recombinant viral vectors were cloned in the genomic DNA backbone of AAV serotype 2, and we generated the hybrid serotype using different capsids from the corresponding serotypes.

Virus injection and head-plate implant. All experimental procedures were approved by the Animal Care and Use Committee at the University of California, Berkeley. For virus injection, young adult (postnatal day (P)40–P60, body weight: 20–30 g) PV-Cre (Jackson Laboratory), CAMKIIα-Cre (Jackson Laboratory), SOM-Cre (Cold Spring Harbor Laboratory) or VIP-Cre (Cold Spring Harbor Laboratory) mice were anaesthetized with ketamine (intraperitoneally; 70 mg per kg body weight) and supplemented with 1.5% isoflurane and 1.5% oxygen. A craniotomy (~0.5 mm diameter) was made above the monocular region of the right V1 (3.7–3.8 mm posterior to bregma, 2 mm lateral from midline), and 1 μl of AAV (containing >1×10^9 viral particles) was injected into the cortex at a depth of 250–500 μm. For recording and behavioural training in awake mice, custom-designed head plates were implanted straight after the virus injection. Small screws (Small Parts) and dental acrylic (Diamond Springs) were used to fix the head plate onto the skull. The skull over the virus-injected area was covered by a silicon sealant (Kwik-Cast, WPI).

Immunohistochemistry. Two weeks after the virus injection, the mice were deeply anaesthetized with isoflurane and immediately perfused with chilled 0.1 M PBS followed by 4% paraformaldehyde (w/v) in PBS. The brain was removed and post-fixed for 15 h at 4 °C. After fixation, the brain was placed in 30% sucrose (w/v) in PBS solution overnight at 4 °C. After embedding and freezing, the brain was sectioned into 40 μm coronal slices using a cryostat (Thermo Fisher). Slices were incubated with blocking solution (2% normal goat serum in PBS with 0.5% Triton X-100) for 2 h at 20 °C and then with primary antibodies diluted in blocking solution overnight at 4 °C. The following primary antibodies were used: anti-PV primary antibody (PVG-214, Swant; 1:1,000), anti-SOM antibody (MAB353, Millipore/Chemicon; 1:200), anti-VIP antibody (20077, ImmunoStar; 1:500) or anti-CAMKIIα antibody (sc-13141, Santa Cruz Biotechnology; 1:50). Slices were then washed three times with the blocking solution and incubated with the secondary antibody for 2 h at 20 °C (for PV and VIP, Alexa594-conjugated anti-rabbit IgG, Invitrogen, 1:1,000; for CAMKIIα, Alexa568-conjugated anti-mouse IgG, Invitrogen, 1:200 and for SOM, Cy3-conjugated anti-rat IgG, Jackson ImmunoResearch, 1:200). Slices were washed three times with PBS (10 min each) and mounted with 4′,6-diamidino-2-phenylindole (DAPI)-containing Vectashield (Vector Laboratories). Fluorescence images were taken under a confocal microscope (Zeiss).

Electrophysiology. Recording experiments were performed 2–6 weeks after virus injection. For anaesthetized experiments, mice were anaesthetized with urethane (intraperitoneally, 1.65 g per kg body weight) supplemented with 1–1.5% isoflurane and restrained in a stereotaxic apparatus (David Kopf Instruments). Body temperature was maintained at 37.5 °C through a heating pad (Harvard Apparatus). For recording in awake mice, the body of the mouse was placed in an acrylic tube (2.9 cm inner diameter; McMaster) and the head plate was fixed on a holder attached to the air table. The mouse could move its body inside the tube while the head was fixed. While the animal was under gas anaesthesia (1.5% isoflurane in oxygen), a craniotomy (~1 mm diameter) was made again above the virus-injected area and a small portion of the dura was removed to allow insertion of a silicon probe (type I, 27 active channels separated by 50 μm, NeuroNexus Technologies; type II, 16 active channels separated by 35 μm, California Probe[6]). Signals were recorded with the Cheetah 32 channel acquisition system (Neuralynx), filtered at 0.6–6 kHz and sampled at 30 kHz. After the experiment, the mouse was euthanized with an overdose of isoflurane. The numbers of mice used for the electrophysiology experiments were 25 anaesthetized PV-ChR2 mice, 33 anaesthetized CAMKIIα-Arch mice, 25 anesthetized SOM-ChR2 mice, 9 anaesthetized VIP-ChR2 mice and 13 awake PV-ChR2 mice.

Whole-cell recordings were made with an Axopatch 700B amplifier (Axon Instruments). Patch pipettes (3–5 MΩ) were filled with internal solution containing (in mM) K-gluconate 125, KCl 5, phosphocreatine 10, MgATP 4, GTP 0.4, HEPES buffer 10 and EGTA buffer 1. Data were filtered at 2 kHz, sampled at 10 kHz and digitized by Digidata 1440 (Molecular Devices), and analysed with custom software in MATLAB (Mathworks). Recordings were made under current

clamp. Step currents (duration, 2 s; amplitude, 0–1.5 nA) were injected with both laser on and off, each repeated two to four times. On the basis of the *F–I* curve (firing rate versus current amplitude; Supplementary Fig. 6), the current threshold (minimal current to evoke spiking) and the slope were determined by linear regression of the curve from the point of initial spiking.

Visual stimulation. Visual stimuli were generated with a PC computer containing a NVIDIA GeForce 6600 graphics board and presented with a XENARC 700V LCD monitor (19.7 cm × 12.1 cm, 960 × 600 pixels, 75 Hz refresh rate, 300 cd m^{-2} maximum luminance, gamma corrected with custom software) located 14 cm from the left eye, positioned such that the receptive fields of the recorded neurons were at the centre of the monitor. To determine the laminar position of each channel of the silicon probe, contrast-reversal checkerboard stimuli were presented at 2 Hz 400 times. For measuring orientation tuning and direction selectivity of V1 neurons, full-field drifting gratings (100% contrast, 2 Hz, 0.04 cycles per degree, 4 s) were presented at 24 directions (separated by 15°) in a pseudorandom sequence. After one block of 24 drifting gratings, 4 s of blank stimulus (grey screen) was presented to measure spontaneous firing rate. A total of eight blocks were presented in each experiment. To measure orientation discrimination in the behavioural experiment, drifting gratings of the same contrast and spatio–temporal frequencies (100% contrast, 2 Hz, 0.04 cycles per degree, 4 s per trial) were presented to the left eye (see below).

Optical activation and silencing. A blue laser (473 nm) combined with a yellow laser (593 nm; CrystaLaser) or a green laser (532 nm; Shanghai Laser & Optics Century Co.) was connected to an output optic fibre and turned on and off by a stimulator (Grass) under computer control. Optical activation of ChR2 was induced by blue light and optical silencing by Arch activation was induced by yellow or green light, focused on top of the craniotomy made for virus injection (for the behavioural experiment, no new craniotomy was made for optical activation). For each PV-ChR2 or CAMKIIα-Arch experiment, the laser power was adjusted manually such that, although the light caused a clear reduction in cortical firing rate, most of the recorded neurons still showed visually driven spiking responses (PV-ChR2, 0.5–1 mW of blue light, with most experiments at 0.6–0.7 mW; CAMKIIα-Arch, 2–15 mW of yellow/green light). For SOM-ChR2 and VIP-ChR2 experiments, we used blue light at powers similar to or higher than in the PV-ChR2 experiment (0.6–1.5 mW). For VIP-ChR2 experiments, increasing the laser power to 2–3 mW also did not cause strong suppression of cortical activity. For measuring changes in spontaneous firing rate induced by optical activation or silencing, 30 trials of light stimulation (5 s per trial, 25 s inter-trial interval) were applied while a blank grey screen was presented as the visual stimulus. To measure the effect of optical activation or silencing on visually evoked responses, laser stimulation began 0.5 s before the onset and ended 0.5 s after the termination of each trial of drifting grating stimulation. Because stray laser light could potentially activate the retina (either from the outside or travelling through the cortex), beginning laser stimulation before the grating stimuli helped to minimize contamination of the responses to the grating stimuli by transient spiking evoked by the onset of stray laser. We alternated between the blocks of trials (24 orientations per block) with and without laser stimulation. For the whole-cell recording experiments (Supplementary Fig. 6), laser stimulation began 200 ms before the onset of current injection and ended 800 ms after termination of the current step. For the behavioural task, laser stimulation started with the trial and lasted for 5 s, thus covering the entire duration of visual stimulation in each trial (Fig. 4b).

Data analysis. To determine the laminar position of each recording channel, multi-unit activity was aligned to the start time of the flash checkerboard stimuli and averaged across trials. Layers four and six were identified on the basis of short-onset latency of the responses. For single-unit isolation, all channels of the silicon probe were separated into groups (four channels per group), and spike waveforms were sorted using Klusters (http://klusters.sourceforge.net)[32]. To assess the quality of each sorted unit, we computed both Lratio and isolation distance[33,34]. To select high-quality single units, we set the thresholds at Lratio <0.1 and Isolation Distance >20 (a unit had to satisfy both criteria to be selected), which corresponded to a <1% error rate[34]. Among these high-quality single units, only those with firing rates greater than 0.1 spikes s^{-1} were included in further analyses. All of the analysis was performed in MATLAB.

To quantify orientation tuning and direction selectivity of each neuron, we fitted the firing rate as a function of orientation by the sum of two Gaussian functions with peaks 180° apart:

$$R(\theta) = a_0 + a_1 e^{\frac{-(\theta-\theta_0)^2}{2\sigma^2}} + a_2 e^{\frac{-(\theta-\theta_0+180^\circ)^2}{2\sigma^2}}$$

in which $R(\theta)$ is the response at orientation θ, a_0 is the untuned component of the response, a_1 and a_2 are the amplitudes of the two Gaussians, θ_0 is the preferred orientation and σ is the standard deviation of the Gaussian function

(Supplementary Fig. 2). Tuning width is measured by σ and the DSI is measured by:

$$|a_1 - a_2|/(a_1 + a_2)$$

The fitting error was computed as:

$$E = \sum_\theta (R_{\text{measure}}(\theta) - R_{\text{fit}}(\theta))^2 / \sum_\theta (R_{\text{measure}}(\theta) - \overline{R})^2$$

in which $R_{\text{measure}}(\theta)$ and $R_{\text{fit}}(\theta)$ are the measured and fitted responses at θ, respectively, and $\overline{R}$ is the measured response averaged across all orientations. We set a threshold of 0.5 for the fitting error; σ and DSI of a cell were included in the analysis only if the fitting error was below the threshold both with and without laser stimulation.

To quantify the reliability of each measured tuning curve, we computed the signal-to-noise ratio (SNR):

$$\text{SNR} = \sum_\theta (R(\theta) - \overline{R})^2 / \sum_\theta \text{VAR}(\theta)$$

in which $R(\theta)$ is the firing rate at θ averaged across all trials, $\overline{R}$ is the mean firing rate across all orientations and $\text{VAR}(\theta)$ is the variance of the response at θ across trials.

The Pearson correlation coefficient (r) was used to quantify the relationship between the laser-induced firing-rate change and changes in tuning width and direction selectivity.

Behavioural experiment. Adult PV-ChR2 mice (P60–P100) were water-deprived for 1 day before starting the training. A total of 35 mice were used for the behavioural experiment. During daily training, the mouse was head-fixed and sat in an acrylic tube within a sound-proof training box. Tongue licks were detected by a custom-made beam-break lickometer. Training ended when the mouse seemed satiated and stopped licking for several minutes[35]. The entire behavioural experiment consisted of five phases: habituation, conditioning, easy discrimination, hard discrimination and optogenetic experiment.

For habituation (2–3 days), there was no visual stimulus and the mouse was given free water rewards (~4 µl) for each lick.

For conditioning (2–3 days), the mouse was trained to lick in response to a visual stimulus (vertically oriented grating drifting rightward; 'go stimulus'). Each trial started with a tone (0.1 s duration; 5 kHz), followed by the visual stimulus (starting 1 s after the tone; 4 s in duration) and ended with an inter-trial period of 4 s. If a lick was detected during the last 2 s of the visual stimulation (response window), the mouse was rewarded with ~4 µl of water for 2 s (hit). If no lick was detected during the response window (miss), water reward was given at the end of the visual stimulus during this conditioning phase. Once the number of hits exceeded 150 within 30 min, the mouse was advanced to the next phase.

For easy discrimination (5–10 days), each trial had the same temporal structure, but the visual stimulus was either the go or no-go (horizontally oriented grating drifting upward) stimulus. The go and no-go trials were randomly interleaved, but the same visual stimulus was never presented more than three consecutive times. Licking within the response window of a go trial (hit) was rewarded with water, whereas licking in the response window of a no-go trial (FA) was punished with a mild air puff (100 ms) and a longer inter-trial interval (8 s, timeout). The mouse was neither rewarded nor punished for a miss (no-lick in a no trial) or correct rejection (CR, no-lick in a no-go trial). Hit and FA rates were quantified as follows:
Hit rate = number of hits/(number of hits + number of misses) and
FA rate = number of FAs/(number of FAs + number of CRs). On the basis of the hit and FA rates, d' was quantified by: $d' = \text{norminv}(\text{hit rate}) - \text{norminv}(\text{FA rate})$ in which norminv is the inverse of the cumulative normal function[20,36]. Higher d' values indicate better performance in visual discrimination.

If a threshold performance ($d' > 1$) was reached within 10 days for the easy discrimination task (orientation difference between go and no-go stimuli,

$\Delta\theta = 90°$), the mouse was advanced to the next phase. Some of the mice failed to reach the threshold, and they were not further tested.

For hard discrimination (>5 days), each test block consisted of 20 trials, in which go and no-go trials at a fixed $\Delta\theta$ were randomly interleaved ($\Delta\theta = 30°$, $60°$ or $90°$). To ensure that the mouse stayed motivated, we alternated between a 'relearning' block ($\Delta\theta = 90°$) and a 'test' block ($\Delta\theta = 30°$, $60°$ or $90°$; the sequence among test blocks of different $\Delta\theta$ was pseudorandom). For each mouse, the orientation of either go or no-go stimulus was fixed, whereas $\Delta\theta$ changed across blocks. d' was then measured as a function of $\Delta\theta$. When $d' > 0.5$ for $\Delta\theta = 30°$, we added test blocks with $\Delta\theta = 10°$ (that is, we alternated between the relearning block with $\Delta\theta = 90°$ and a test block with $\Delta\theta = 10°$, $30°$, $60°$ or $90°$). The addition of this most difficult block ($\Delta\theta = 10°$) often caused a considerable drop in the overall performance, probably caused by a loss of motivation owing to the high failure rate. We continued the training in this phase until the performance recovered to a level comparable to that before adding the block with $\Delta\theta = 10°$.

For the optogenetic experiment, the effect of optical activation of PV$^+$ neurons on orientation discrimination was measured at $\Delta\theta = 10°$, $30°$, $60°$ and $90°$. In each block, laser stimulation was applied in 50% of randomly selected trials, and d' was analysed separately for trials with and without laser stimulation. To eliminate the trials near the end of each session, when the mouse was satiated and lost motivation for the task, we only included the trials with >50% hit rate within 100 consecutive trials. To minimize the amount of laser light reaching the eyes, the optic fibre was shielded by a black tape.

Because very strong suppression of visually driven cortical responses can negatively affect perception, the laser power was chosen carefully to match between the behavioural and electrophysiological experiments. We first measured the loss of laser power through the skull (craniotomy was made for the acute electrophysiological experiments but not for the chronic behavioural experiments) in vitro by passing the laser beam through the skull of a mouse previously used in the behaviour experiment (note that a small hole was drilled during virus injection several weeks before the behavioural experiment, so the skull at the injection site is thinner than at other places; our calibration was performed at the injection site). We found ~30% loss of power after passing through the skull. Because in most of the PV-ChR2 electrophysiology experiments we used 0.6–0.7 mW of laser power, which was found to be effective in sharpening the tuning curve without excessive suppression of cortical activity, we chose 0.8–1 mW laser power for the behavioural experiments. Each mouse was tested in up to eight sessions (one session per day), and all animals tested with the 0.8–1 mW laser were included ($n = 25$).

Among the 25 PV-ChR2 mice tested, 6 were advanced directly from easy discrimination to optogenetic experiment, and the effect of PV activation was tested only at $\Delta\theta = 90°$; 10 mice were trained in hard discrimination at $\Delta\theta = 30°$, $60°$ and $90°$ and the effect of laser was tested at these three angles; the remaining 9 mice were trained in hard discrimination at $\Delta\theta = 10°$, $30°$, $60°$ and $90°$ before the effect of laser was tested at all four angles.

31. Maheshri, N. et al. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nature Biotechnol.* **24,** 198–204 (2006).
32. Hazan, L. et al. Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. *J. Neurosci. Methods* **155,** 207–216 (2006).
33. Harris, K. D. et al. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* **84,** 401–414 (2000).
34. Schmitzer-Torbert, N. et al. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131,** 1–11 (2005).
35. O'Connor, D. H. et al. Vibrissa-based object localization in head-fixed mice. *J. Neurosci.* **30,** 1947–1967 (2010).
36. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, 1966).

LETTER

Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells

Ping Wei[1,2]*, Wilson W. Wong[1,2,3]*†, Jason S. Park[1,2,3], Ethan E. Corcoran[2,3,4]†, Sergio G. Peisajovich[1,2]†, James J. Onuffer[1,3], Arthur Weiss[2,3,4] & Wendell A. Lim[1,2,3]

Bacterial pathogens have evolved specific effector proteins that, by interfacing with host kinase signalling pathways, provide a mechanism to evade immune responses during infection[1,2]. Although these effectors contribute to pathogen virulence, we realized that they might also serve as valuable synthetic biology reagents for engineering cellular behaviour. Here we exploit two effector proteins, the *Shigella flexneri* OspF protein[3] and *Yersinia pestis* YopH protein[4], to rewire kinase-mediated responses systematically both in yeast and mammalian immune cells. Bacterial effector proteins can be directed to inhibit specific mitogen-activated protein kinase pathways selectively in yeast by artificially targeting them to pathway-specific complexes. Moreover, we show that unique properties of the effectors generate new pathway behaviours: OspF, which irreversibly inactivates mitogen-activated protein kinases[4], was used to construct a synthetic feedback circuit that shows novel frequency-dependent input filtering. Finally, we show that effectors can be used in T cells, either as feedback modulators to tune the T-cell response amplitude precisely, or as an inducible pause switch that can temporarily disable T-cell activation. These studies demonstrate how pathogens could provide a rich toolkit of parts to engineer cells for therapeutic or biotechnological applications.

Many bacterial pathogens have developed an array of effector proteins to rewire host signalling networks and downregulate the immune response[2] (Fig. 1a). Some effectors mimic host activities, such as the *Y. pestis* effector YopH, which is a highly active phosphotyrosine phosphatase[3]. Other effectors use unusual mechanisms, such as the *S. flexneri* OspF protein, which irreversibly inactivates mitogen-activated protein kinases (MAPKs) by catalysing a β-elimination reaction that removes the hydroxyl group of the key phospho-threonine side chain[4].

MAPK pathways have a central role in diverse eukaryotic responses, ranging from immune response to cell-fate decisions[5,6]. Thus, the ability to tune MAPK response would facilitate engineering cells for diverse therapeutic and biotechnological applications[7,8]. Recent work has shown that MAPK signalling dynamics in yeast can be reshaped with synthetic feedback loops that involve controlled expression and targeting of pathway modulators to appropriate signalling complexes[9]. Identifying effective pathway modulators is challenging; thus we considered that pathogen effector proteins may have untapped use as components for predictably and systematically engineering signalling pathways. Here we use the effector proteins OspF and YopH to modulate kinase signalling pathways in yeast and in human primary T cells.

We first introduced OspF into yeast. As reported[10], overexpression of OspF led to growth inhibition under standard conditions, hyperosmotic stress conditions (Fig. 1b) and cell-wall-damaging conditions (Supplementary Fig. 1a). OspF contains a canonical docking peptide at its amino (N) terminus that allows it to bind several MAPKs in yeast[11]. We found that expression of an OspF mutant lacking its native docking peptide (ΔN-OspF) yielded normal growth behaviour under all conditions (Fig. 1b and Supplementary Fig. 1a). Next we tested whether ΔN-OspF could be redirected to a specific pathway by tagging the protein with a leucine zipper heterodimerization motif, and fusing the complementary interacting motif to Pbs2, the scaffold protein that organizes the osmolarity MAPK pathway. This targeted version of ΔN-OspF only showed a growth defect under high salt conditions, showing that OspF activity could be engineered to inhibit a specific MAPK (Fig. 1b).

To explore re-targeting OspF to specific pathways further, we engineered yeast strains in which ΔN-OspF was selectively targeted to either the osmolarity MAPK complex or the mating MAPK complex (by targeting it to the mating pathway scaffold protein, Ste5) (Fig. 1c). Targeting of ΔN-OspF to the Pbs2 inhibited the osmolarity response but not the mating response. Conversely, when ΔN-OspF was targeted to Ste5, only the mating response was inhibited. Thus, the inhibitory activity of this effector could be selectively aimed at one of several MAPK pathways in the same cell.

One of the unique aspects of OspF is that it catalyses an irreversible inactivation of MAPKs (unlike reversible dephosphorylation by a phosphatase). Thus MAPK activity can only be restored through new protein synthesis, which has a much slower timescale than re-phosphorylation (Fig. 2a, b). This longer timescale would be expected to lead to an extended refractory period after OspF action, during which the targeted MAPK pathway could no longer respond to subsequent stimuli.

Computational simulations indicated that a long refractory period could result in significant changes to the frequency-dependent behaviour of pathway response (Supplementary Fig. 2). There is growing evidence that cells use frequency modulation of diverse molecular events to encode and transmit information[12,13]. Our models indicated that with a negative feedback loop (that is, MAPK activity induced expression of OspF), pathway output would be dampened when input periods are long enough to accumulate significant amounts of the negative effector but shorter than the refractory period (Supplementary Fig. 2).

To test if OspF could be used to filter frequency dependent inputs, we constructed a synthetic negative feedback loop in the yeast osmo-response pathway by expressing OspF targeted to the osmo-response signalling complex (ΔN-OspF–zipper) from the Hog1 responsive promoter, pSTL1 (Fig. 2c). As a comparison, we also engineered an analogous synthetic feedback loop using a reversible Hog1 MAPK inhibitory protein—the yeast MAPK phosphatase, PTP2. Phospho-Hog1 translocation to the nucleus was used as a fast-timescale output reporter (Supplementary Fig. 3)[14–16]. To measure integrated output over a longer timescale, we also measured a slower timescale

[1]Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, California 94158, USA. [2]Howard Hughes Medical Institute, University of California San Francisco, San Francisco, California 94158, USA. [3]The Cell Propulsion Laboratory, University of California San Francisco, San Francisco, California 94158, USA. [4]Department of Medicine, University of California San Francisco, San Francisco, California 94158, USA. †Present addresses: Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA (W.W.W.); Department of Pulmonology and Critical Care Medicine, Northwest Permanente P.C., 10180 SE Sunnyside Road, Clackamas, Oregon 97015, USA (E.E.C.); Department of Cell & Systems Biology, University of Toronto, Toronto, Ontario M5S 3G5, Canada (S.G.P.).
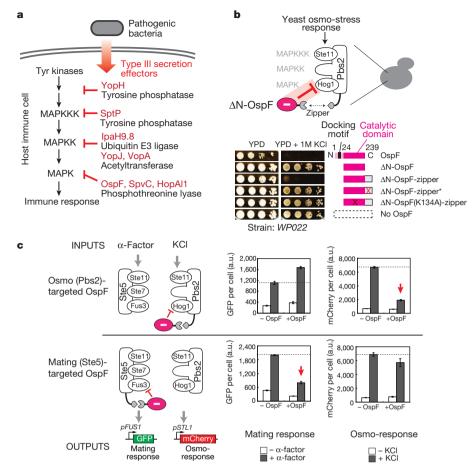*These authors contributed equally to this work.

Figure 1 | Bacterial effector OspF can block selective MAPK pathways in yeast. a, Type III secretion effectors that modulate host kinase signalling. MAPKKK, MAP-kinase kinase kinase. **b**, Targeting of OspF to yeast osmolarity pathway. Wild-type OspF impairs growth on rich media, but is rescued by docking motif deletion (ΔN-OspF). Recruitment of ΔN-OspF to osmolarity

scaffold Pbs2 by leucine zipper selectively blocks growth on 1 M KCl (zipper*, mutant leucine zipper; K134A, catalytic dead mutant of OspF). **c**, ΔN-OspF selectively inhibits mating or osmolarity if targeted to appropriate scaffold complex, assayed using pathway-specific transcriptional reporters. Average fluorescence and s.d. of three experiments are shown. a.u., arbitrary units.
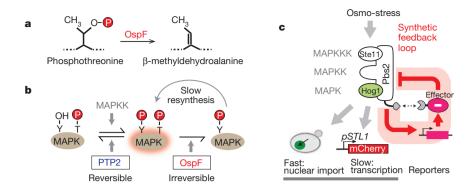
transcriptional reporter—expression of mCherry from the *pSTL1* promoter.

We found that the OspF-mediated negative feedback circuit altered the osmo-stress pathway response to intermediate frequency stimulation, but not to continuous stress or to high-frequency stimulation (Fig. 2d and Supplementary Fig. 4a). Furthermore, in the course of stimulating cells with pulses of KCl of varying length, we discovered that an input period of approximately 16 min (intermediate frequency) leads to highly divergent transcriptional responses (Fig. 2d). Examination of Hog1 nuclear import in the OspF feedback strain shows that after three pulses, the amount of Hog1 competent for nuclear localization has decreased to near zero, consistent with a model in which—in this timeframe—there is sufficient OspF to inactivate the bulk of the Hog1 population and now to render the cells refractory to further pulses of stimulation. The cells containing a PTP2 negative feedback circuit do not show this marked filtering at this frequency. With higher frequency stimulation (2 min period), the three strains also do not significantly differ from each other in response, presumably because there is inadequate activation time with each pulse to build up a sufficient concentration of effector[17].

A broad frequency dependence analysis shows that the wild-type osmo-response pathway functions as a band-pass filter, with maximal response at intermediate frequencies, whereas the engineered pathway more closely resembles a low-pass filter, with maximal responses at low frequency (Fig. 2e). These distinct frequency filtering behaviours fit those predicted by computational simulations (Supplementary Fig. 2d).

We then sought to test whether these bacterial effectors could be used to rewire signalling in immune cells. Human T cells are an attractive synthetic biology platform because they can be isolated from patients, genetically engineered *ex vivo* and then transferred back into patients to treat cancer and chronic infection[18,19]. Although promising, the therapeutic application of engineered T cells carries risks of adverse side effects including inadvertent autoimmune-like attack of off-target host tissues[20,21]. Thus mechanisms to control the specificity, amplitude and timing of T-cell function are critical to balance therapeutic action against off-target toxicity.

Both OspF and YopH can modify the T-cell receptor (TCR) pathway (Fig. 3a). OspF inactivates the MAPK extracellular signal-regulated kinase (ERK), which is a central component of TCR signalling[4,22], whereas YopH dephosphorylates phospho-tyrosine, including the T-cell scaffold proteins LAT and SLP-76 (ref. 23). Constitutive expression of YopH and OspF in *Jurkat* T cells leads to severe inhibition of TCR activation, as measured by an NFAT transcriptional reporter[24] (Fig. 3a) (as well as other reporters of T-cell activation; Supplementary Fig. 5a). Expression of the catalytic dead versions of YopH and OspF had no effect on the TCR activation (Supplementary Fig. 5b). In addition, we showed that these two effectors clearly target distinct steps of the T-cell activation pathway, because induction of the T cells with the combination of phorbol 12-myristate 13-acetate (PMA) and ionomycin, which activates the T cell downstream of the PLCγ1-LAT/SLP76 dependent response, bypassed YopH inhibition[25], but was sensitive to OspF inhibition (Fig. 3a and Supplementary Fig. 5a). Thus, distinct pathogen effector proteins can be used to block this
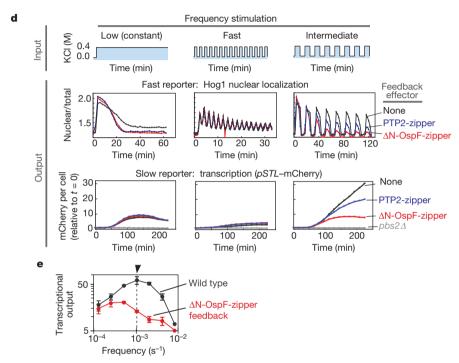
Figure 2 | Tuning-frequency-dependent response of yeast osmolarity pathway using synthetic OspF feedback loop. a, OspF is MAPK phosphothreonine lyase. **b**, OspF irreversibly inhibits MAPK activity. **c**, Synthetic negative feedback loop was built by expressing OspF from osmo-inducible *pSTL1* promoter. Hog1–GFP nuclear accumulation (Nuc:Tot ratio) was measured as a fast output reporter. *pSTL1*–mCherry served as a slower, gene expression reporter. **d**, Cells were stimulated with variable frequency

inputs: low (constant), high (period = 2 min) and intermediate (period = 16 min). Significant difference in response to intermediate frequency stimulation is observed. **e**, Frequency–response curves for wild-type and OspF feedback strains (arrow: period = 16 min). Each point is average of 50–100 cells; s.d. from three repeats is shown. More detail on the frequency analysis is given in Supplementary Figs 2–4.

pathway at particular steps, much like a specific small-molecule inhibitor.

Given the ability of OspF and YopH to modulate T-cell responses, we sought to use these proteins to build circuits that could, in principle, improve the safety of therapeutic T cells. In adoptive T-cell therapy, a challenge is to limit over-activation or off-target activation of T cells that could lead to killing of host cells or to cytokine storm—a life-threatening immune response. One approach is to incorporate a safety 'kill switch'[26–28] into the T cells, such as the herpes simplex virus thymidine kinase gene. This protein converts the pro-drug ganciclovir into an inhibitor of replication, thus killing cells expressing the gene. Although herpes simplex virus thymidine kinase is currently being tested in a phase III clinical trial for the treatment of graft versus host disease in bone marrow transplants, this strategy irreversibly destroys the engineered, adoptively transferred cells[29]. Thus, instead of killing the engineered cells, we sought to design circuits that would limit the amplitude of the T-cell response or to pause T-cell activity temporarily.

We first tested whether bacterial effectors could be used to limit the response amplitude of *Jurkat* T cells. Negative feedback loops can act

to limit the maximal amplitude of a response[30], so we engineered a library of negative feedback loops in which the OspF and YopH were expressed from a series of TCR-responsive promoters of varying strength (AP1 and NFAT) (Fig. 3b and Supplementary Fig. 7). For further tuning of feedback parameters, we also tagged effectors with degradation sequences (PEST motif) that reduce the half-life of the effectors. This series of negative feedback loops led to controlled reduction of the maximal response amplitude of T-cell activation (Fig. 3b). Moreover, the amplitude could be tuned systematically by varying feedback promoter strength and effector stability (Supplementary Fig. 7b). For example, expression of OspF from the strong feedback (*pNFAT*) promoter leads to a very low maximal response amplitude, but this effect could be systematically tuned by destabilizing the OspF effector with a PEST tag.

We also tested whether the bacterial effectors could be used to construct pause switches, which could transiently and reversibly disable T cells. We placed the effectors under the control of a tetracycline inducible promoter (*pTRE*), which allowed external control of the timing of effector expression with the addition of doxycycline.
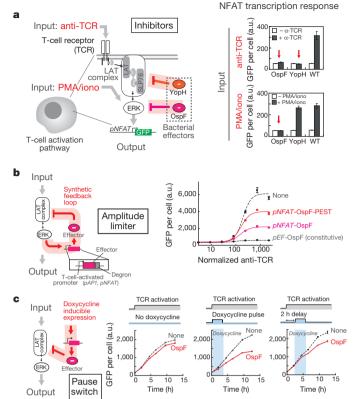
Figure 3 | OspF can be used to control T-cell activation amplitude and duration precisely in *Jurkat* T cells. a, Constitutively expressed OspF and YopH inhibit specific steps in the TCR pathway. Activation of NFAT transcription by anti-TCR antibody versus PMA/ionomycin is shown. **b**, Synthetic amplitude limiters were constructed using negative feedback loops with different effectors, promoters and degrons. Dose–response curves of the *pNFAT*-OspF circuit, with or without degron, are shown. **c**, Synthetic pause switch was built by expressing OspF from a doxycycline-inducible promoter, *pTRE*. A 4 h pulse induction disabled T-cell response when initiated at 0 or 2 h after T-cell stimulation. Standard deviation from three samples is shown.

Effectors were fused to a destabilization domain so that they would be rapidly degraded once doxycycline is removed. Using this system, we first showed that transient expression of bacterial effectors can inhibit TCR signalling after the pathway is activated in *Jurkat* T cells (Fig. 3c). TCR signalling can be inhibited up to 6 h after activation using this approach (Supplementary Fig. 8b). Finally, we showed that engineered T cells can be subjected to cycles of TCR activation, paused with a short period of induced expression of the bacterial effector, and then reactivated after this pause (Supplementary Fig. 8c).

We then tested the pause switch in a clinically important cell type for adoptive immunotherapy—primary human CD4$^+$ T cells (in contrast to the *Jurkat* T-cell line, which does not require cytokine or TCR activation to stimulate proliferation). We showed that when OspF is induced by the addition of doxycycline, both interleukin-2 (IL-2) release and cell proliferation were inhibited in a dose-dependent manner (Fig. 4b, c). Activation of the TCR by anti-CD3/CD28 and antigen-presenting cells can also be inhibited by expression of OspF (Supplementary Fig. 9a). Moreover, after doxycycline is removed, IL-2 release and cell division recovers in 6–18 h (Fig. 4d, e). Sustained exposure to doxycycline can inhibit T-cell activity over the course of several days (Fig. 4e) without having any significant effect on cell viability (Supplementary Fig. 9c). Thus this work provides a proof of principle for the design of a simple 'pause' switch that could allow external control over the timing and level of T-cell activation and cytokine release, to minimize adverse events associated with adoptive immunotherapy such as cytokine storm.
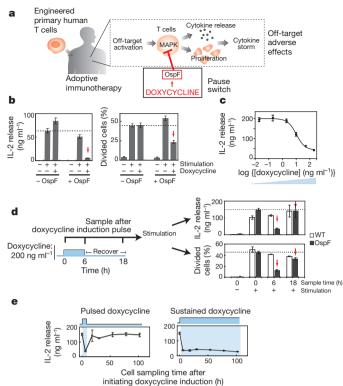

Figure 4 | OspF can be used as a synthetic pause switch to control human primary CD4$^+$ T-cell activation. a, Off-target activation of transplanted T cells can induce cytokine storm in adoptive immunotherapy. A pause switch could prevent this adverse response. **b**, Six-hour pre-induction of OspF by doxycycline inhibited IL-2 release and proliferation of activated CD4$^+$ T cells. **c**, Six-hour pre-treatment with different doses of doxycycline can tune IL-2 release. **d**, Cells treated with 6 h doxycycline pulse were sampled at different times after the pulse, then subjected to a 24 h IL-2 release assay or a 4 day proliferation assay. **e**, IL-2 release can be inhibited either transiently or in a sustained manner by varying the duration of doxycycline treatment. Average and s.d. of three experimental repeats are shown.

Most work on bacterial pathogen effector proteins has the long-term aim of neutralizing the pathogens' infectious capabilities. We have shown, however, that bacterial effectors can also be valuable synthetic biology tools, because of their unique biochemical properties. We have used bacterial effectors to modulate MAPK signalling in yeast to generate novel time-dependent dynamics. We have also shown that bacterial effectors can be used to tune human TCR signalling dynamics flexibly, with potential application as safety switches for adoptive immunotherapy. The vast array of bacterial pathogen effector proteins, beyond those studied here, holds promise as a rich and important source of parts for the cellular engineering toolkit.

METHODS SUMMARY

Flow cytometry experiments. Analysis of the pathway-dependent fluorescent protein (green fluorescent protein (GFP) and mCherry) expression in yeast cells and phosphorylated ERK in *Jurkat* T cells was performed with a BD LSRII flow cytometer (BD Biosciences) equipped with a high-throughput sampler. α-Factor (1.5 μM; GenScript) or 0.4 M of KCl were added into yeast cultures to induce separately the mating- or osmo-specific pathway response. For staining of phosphorylated ERK in *Jurkat* T cells, cells were fixed, made permeable by incubation with ice-cold 90% methanol on ice for 30 min and stained with primary antibody to phosphorylated ERK (Cell Signaling) and anti-rabbit APC secondary antibody (Jackson Immunoresearch).

Microfluidics and fluorescent microscopy. Microfluidic yeast cell culture was performed in Y04C plates with an ONIX flow control system (Cellasic). Cells were loaded into the flow chamber pre-coated with concanavalin A. Pulse stimulation with salt media was performed by ONIX FG flow control software (Cellasic) with a flow pressure of 8 pounds per square inch. Image acquisition was performed with a TE2000-E automated inverted microscope (Nikon) with Perfect Focus and a ×100

oil immersion lens. Image analysis for both nuclear Hog1–GFP import and *pSTL1*–mCherry expression was performed with custom MATLAB (Mathworks) software.

Human T-cell activation assay. Resting human primary CD4$^+$ cells (transduced with the pause switch constructs or untransduced) were pretreated with 200 ng ml^{-1} doxycycline for 6 h. Fifty thousand cells were placed in a 96-well plate with 200 µl human growth media with activation agents added (10 ng ml^{-1} PMA + 0.5 µM ionomycin, magnetic Dynabeads coated with anti-CD3/anti-CD28 (beads:cells ratio, 0.3:1)). After 24 h of incubation at 37 °C, the released IL-2 in the supernatant was measured with a human IL-2 ELISA kit II (BD Biosciences). Cells labelled with CellTrace Violet dye (Invitrogen) were assayed by flow cytometry after incubation for 4 days to quantify cell proliferation.

Full Methods and any associated references are available in the online version of the paper.

1. Ribet, D. & Cossart, P. Pathogen-mediated posttranslational modifications: a re-emerging field. *Cell* **143**, 694–702 (2010).
2. Broberg, C. A. & Orth, K. Tipping the balance by manipulating post-translational modifications. *Curr. Opin. Microbiol.* **13**, 34–40 (2010).
3. Zhang, Z. Y. *et al.* Expression, purification, and physicochemical characterization of a recombinant *Yersinia* protein tyrosine phosphatase. *J. Biol. Chem.* **267**, 23759–23766 (1992).
4. Li, H. *et al.* The phosphothreonine lyase activity of a bacterial type III effector family. *Science* **315**, 1000–1003 (2007).
5. Shaw, A. S. & Filbert, E. L. Scaffold proteins and immune-cell signalling. *Nature Rev. Immunol.* **9**, 47–56 (2009).
6. Dong, C., Davis, R. J. & Flavell, R. A. MAP kinases in the immune response. *Annu. Rev. Immunol.* **20**, 55–72 (2002).
7. Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nature Rev. Genet.* **11**, 367–379 (2010).
8. Purnick, P. E. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nature Rev. Mol. Cell Biol.* **10**, 410–422 (2009).
9. Bashor, C. J., Helman, N. C., Yan, S. & Lim, W. A. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* **319**, 1539–1543 (2008).
10. Kramer, R. W. *et al.* Yeast functional genomic screens lead to identification of a role for a bacterial effector in innate immunity regulation. *PLoS Pathog.* **3**, e21 (2007).
11. Zhu, Y. *et al.* Structural insights into the enzymatic mechanism of the pathogenic MAPK phosphothreonine lyase. *Mol. Cell* **28**, 899–913 (2007).
12. Nelson, D. E. *et al.* Oscillations in NF-κB signaling control the dynamics of gene expression. *Science* **306**, 704–708 (2004).
13. Cai, L., Dalal, C. K. & Elowitz, M. B. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* **455**, 485–490 (2008).
14. Muzzey, D., Gomez-Uribe, C. A., Mettetal, J. T. & van Oudenaarden, A. A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell* **138**, 160–171 (2009).
15. Mettetal, J. T., Muzzey, D., Gomez-Uribe, C. & van Oudenaarden, A. The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. *Science* **319**, 482–484 (2008).
16. Hersen, P., McClean, M. N., Mahadevan, L. & Ramanathan, S. Signal processing by the HOG MAP kinase pathway. *Proc. Natl Acad. Sci. USA* **105**, 7165–7170 (2008).
17. Pelet, S. *et al.* Transient activation of the HOG MAPK pathway regulates bimodal gene expression. *Science* **332**, 732–735 (2011).
18. Morgan, R. A., Dudley, M. E. & Rosenberg, S. A. Adoptive cell therapy: genetic modification to redirect effector cell specificity. *Cancer J.* **16**, 336–341 (2010).
19. June, C. H., Blazar, B. R. & Riley, J. L. Engineering lymphocyte subsets: tools, trials and tribulations. *Nature Rev. Immunol.* **9**, 704–716 (2009).
20. Morgan, R. A. *et al.* Case report of a serious adverse event following the administration of T cells transduced with a chimeric antigen receptor recognizing ERBB2. *Mol. Ther.* **18**, 843–851 (2010).
21. Brentjens, R., Yeh, R., Bernal, Y., Riviere, I. & Sadelain, M. Treatment of chronic lymphocytic leukemia with genetically targeted autologous T cells: case report of an unforeseen adverse event in a phase I clinical trial. *Mol. Ther.* **18**, 666–668 (2010).
22. Arbibe, L. *et al.* An injected bacterial effector targets chromatin access for transcription factor NF-κB to alter transcription of host genes involved in immune responses. *Nature Immunol.* **8**, 47–56 (2007).
23. Gerke, C., Falkow, S. & Chien, Y. H. The adaptor molecules LAT and SLP-76 are specifically targeted by *Yersinia* to inhibit T cell activation. *J. Exp. Med.* **201**, 361–371 (2005).
24. Lin, J. & Weiss, A. The tyrosine phosphatase CD148 is excluded from the immunologic synapse and down-regulates prolonged T cell signaling. *J. Cell Biol.* **162**, 673–682 (2003).
25. Yao, T., Mecsas, J., Healy, J. I., Falkow, S. & Chien, Y. Suppression of T and B lymphocyte activation by a *Yersinia pseudotuberculosis* virulence factor, yopH. *J. Exp. Med.* **190**, 1343–1350 (1999).
26. Kieback, E., Charo, J., Sommermeyer, D., Blankenstein, T. & Uckert, W. A safeguard eliminates T cell receptor gene-modified autoreactive T cells after adoptive transfer. *Proc. Natl Acad. Sci. USA* **105**, 623–628 (2008).
27. de Witte, M. A. *et al.* An inducible caspase 9 safety switch can halt cell therapy-induced autoimmune disease. *J. Immunol.* **180**, 6365–6373 (2008).
28. Bonini, C. *et al.* The suicide gene therapy challenge: how to improve a successful gene therapy approach. *Mol. Ther.* **15**, 1248–1252 (2007).
29. Ciceri, F. *et al.* Infusion of suicide-gene-engineered donor lymphocytes after family haploidentical haemopoietic stem-cell transplantation for leukaemia (the TK007 trial): a non-randomised phase I–II study. *Lancet Oncol.* **10**, 489–500 (2009).
30. Brandman, O. *et al.* Feedback loops shape cellular signals in space and time. *Science* **322**, 390–395 (2008).

Author Contributions P.W., S.G.P. and W.A.L. initiated the project in yeast. W.W.W., E.E.C., A.W. and W.A.L. initiated the project in T cells. P.W., W.W.W. and W.A.L. wrote the paper. P.W. and S.G.P. planned and performed the experiments in yeast. W.W.W., P.W., E.E.C., J.J.O. and J.S.P. planned and performed the experiment in T cells.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to W.A.L. (lim@cmp.ucsf.edu).

METHODS

Yeast constructs and strains. All yeast constructs used in this study (Supplementary Table 1) were cloned using a combinatorial cloning strategy on the basis of the type II restriction enzyme AarI developed by Peisajovich et al.[31]. For the MAPK pathway specific inhibitory experiments, strains WP022 and WP116 were constructed from the W303-derived strain SP147 (Supplementary Table 2) by tagging the carboxy (C) terminus of endogenous Pbs2 or Ste5 with leucine zipper, respectively. To construct WP039 for measuring the nuclear Hog1 accumulation, the C terminus of endogenous Hog1 in wild-type W303 strain was tagged with GFP using standard integration technique. For the visualization of the nucleus, a histone protein Htb2 was C-terminally tagged with mCherry, expressed from the ADH1 promoter, and integrated at TRP1 locus. Endogenous Pbs2 was tagged with a leucine zipper by PCR integration and was used in synthetic feedback loops. Synthetic effector gene cassettes were integrated at the LEU2 locus. Yeast promoters, terminator and the PTP2 gene were PCR amplified from Saccharomyces cerevisiae genomic DNA (Invitrogen). YopH gene was a gift from K. Orth. The OspF gene was synthesized by Integrated DNA Technologies. All yeast genomic integrations were confirmed by yeast colony PCR.

Yeast flow cytometry experiments. Analysis of the pathway-dependent fluorescent protein (GFP and mCherry) expression in yeast cells was performed with the BD LSRII flow cytometer (BD Biosciences) equipped with a high-throughput sampler. For each experiment, triplicate cultures were grown in synthetic complete media to early log phase ($D_{600\,nm} = 0.1$–0.2). At time = 0, 1.5 μM of α-factor (GenScript) or 0.4 M of KCl were added into parallel cultures to induce separately the mating-specific or osmo-specific pathway response. For mating response, 100 μl aliquots were taken at time = 0 and after 2 h of induction; for osmotic response, 100 μl aliquots were taken at time = 0 and after 1 h of induction. Each 100 μl sample aliquot was immediately mixed with 100 μl of cycloheximide (10 μg ml^{-1}) in 96-well plates to stop the protein synthesis. After incubating the samples at room temperature for 30 min in the dark to allow for the maturation of fluorescent proteins, the levels of fluorescence protein were determined by flow cytometry. For each read, 10,000 cells were counted, and the mean fluorescent density was calculated by Flowjo software (BD Biosciences) as the pathway output value, and the s.d. from triplicate experiments was indicated as the error bar.

Yeast microfluidics, fluorescent microscopy and image processing. Microfluidic yeast cell culture was performed in Y04C plates with an ONIX flow control system (Cellasic)[32]. Cultures were grown to mid-log phase ($D_{600\,nm} = 0.2$–0.8) in synthetic complete media. Cells were diluted to $D_{600\,nm} = 0.01$ in 600 μl of fresh media and sonicated at a minimum set of 11% for 1 s using a Fisher Scientific model 500 sonicator with a 2 mm tip. Cells were loaded into the flow chamber pre-coated with concanavalin A and flowed over by synthetic complete media for more than 20 min before applying the square pulse sequence. Pulse stimulation with salt media was performed by ONIX FG flow control software (Cellasic) with the flow pressure of 8 pounds per square inch. Image acquisition was performed with a TE2000-E automated inverted microscope (Nikon) with Perfect Focus and ×100 oil immersion lens.

Background subtraction was performed first on all fluorescence images using ImageJ (http://imagej.nih.gov). The subsequent image analysis for both nuclear Hog1–GFP import and pSTL1–mCherry expression was performed with custom MATLAB (Mathworks) software (developed by K.-Y. Lau from C. Tang's laboratory at University of California San Francisco). For Hog1–GFP import analysis, cell boundaries were first determined from the bright-field differential interference contrast microscopy images. Cell nuclei were segmented by mCherry-labelled nuclear images. The nuclear and total cell GFP densities were calculated from the GFP fluorescent images. For pSTL1–mCherry expression, cell boundaries were determined the same way Hog1–GFP localization and mCherry densities were determined from the mCherry fluorescent images. For all frequency responses, a 210 min time course of pulse stimulation was performed and the maximum value of each time course was taken as the output of the stimulation frequency. The population average of more than 50–100 cells was determined for each single measurement, each experiment was repeated at least three times (see also Supplementary Fig. 3c, d) and the s.d. was calculated with the three repeats.

Jurkat T-cell lines, plasmids and transfection. The Jurkat T cell with pNFAT–EGFP (neomycin resistant) stably integrated into the chromosome is a Weiss laboratory stock strain. All plasmids were made by using standard cloning techniques, AarI combinatorial cloning technique[31] and Gateway cloning technique (Invitrogen). See Supplementary Tables for more cloning detail. Jurkat cells were maintained in RPMI 1640 supplemented with 10% heat-inactivated FBS (Invitrogen), extra L-glutamine (2 mM), penicillin, streptomycin and G418 (2 mg ml^{-1}, Invitrogen).

For transfection, Jurkat T cells were cultured in RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum (FBS) and glutamine (10G RPMI) for at least 1 day. Twenty million cells were spun down, washed once with 10G RPMI

and re-suspended in 300 μl of 10G RPMI. Fifteen micrograms of each plasmid were added to each transfection, vortexed briefly and incubated at room temperature for 15 min. The cell/DNA mixture was then subject to electroporation (BioRad, square pulse, 300 V, 10 ms pulse, 0.4 cm cuvette). The cells were rested at room temperature for 10 min before re-suspending in 10 ml of 10G RPMI. The cells were allowed to recover overnight before performing further experiments. For transfection with plasmids that contained the pTRE promoter, the serum was switched to Tet-free serum (Clontech).

T-cell activation and doxycycline induction. TCR was activated with a Jurkat-specific anti-TCR antibody, C305 (ascites, University of California San Francisco antibody core) at 1:2,000 dilution unless stated otherwise at a cell density of 2.5 million cells per millilitre (for most experiments) or 0.5 million cells per millilitre (for doxycycline-inducible expression (100 nM) of effectors experiment). Phorbol 12-myristate 13-acetate (25 ng ml^{-1}) and ionomycin (1 μM) was also used to activate T cells in some experiment. For the dose–response curve, the highest dose of C305 was 1:600, serially diluted threefold to generate eight doses in total.

T-cell antibody staining and flow cytometry analysis. For staining of cell surface expression of CD69, cells were fixed and stained with anti-CD69-APC (BD). For staining of phosphorylated ERK, cells were fixed, made permeable by incubation with ice-cold 90% methanol on ice for 30 min and stained with primary antibody to phosphorylated ERK (4370 Cell Signaling) and anti-rabbit APC secondary antibody (711-136-152, Jackson Immunoresearch). All samples were analysed with a BD LSRII equipped with a high-throughput sampler. Live cells were determined from forward and side scattering. Transfected cells were determined by comparing cells without mCherry to cells transfected with pEF–mCherry. Only transfected and live cells were included in the analysis. Error bar represents the s.d. from three samples.

Western blot to determine protein expression level. One million live cells were quickly spun down and lysed on ice for 30 min. The supernatant was then spun down at 4 °C for 30 min. The lysates were then mixed with DTT and SDS sample loading buffer and boiled for 3 min. Samples were separated with SDS–polyacrylamide gel electrophoresis gel (4–12% Bis–Tris) and then transferred to nitrocellulose blot. The blot were stained with primary anti-HA antibody (Santa Cruz Biotechnology) and Li-Cor anti-mouse 680 LT secondary antibody. The blot was imaged with a Li-Cor Odyssey.

Human primary CD4$^+$ T-cell transduction. Human peripheral blood mononuclear cells were collected from normal donors and acquired as cell suspensions from flushed TRIMA leukoreduction chambers (Blood Centers of the Pacific). Primary CD4$^+$ T cells were purified by negative selection and Ficoll-Paque PLUS density medium separation (RosetteSep, Stem Cell Technologies). Purified cells were cryopreserved and placed in liquid nitrogen storage.

Replication-incompetent lentiviral particles were prepared in 293T cells by standard methods. Briefly, constructs of interest were cloned into the transfer vector pHR′SIN:CSW using standard molecular biology techniques and then co-transfected into 293T cells with the viral packaging plasmids pCMVdR8.91 and pMD2.G using the transfection reagent FuGENE HD (Roche). Amphotropic VSV-G pseudotyped lentiviral particles in the supernatant were collected 48 h later.

Before use, human primary CD4$^+$ T cells were thawed and cultured overnight in growth medium (X-VIVO 15 + 5% human AB serum + 10 mM N-acetylcysteine + 1× β-mercaptoethanol + 1× Primocin) supplemented with 30 U ml^{-1} IL-2. The next day, cells were activated with Dynabeads human T-Activator CD3/CD28 (Invitrogen) at a 3:1 beads-to-cells ratio. After 24 h of activation, the cells were transduced with lentiviral particles. In some cases, transduction was performed on RetroNectin-coated tissue culture plates to enhance viability and transduction efficiency. Briefly, non-tissue-culture-treated plates were coated with RetroNectin (32 μg ml^{-1}) and then blocked with PBS + 2% BSA. Viral supernatant was loaded into the wells and the plate was centrifuged at 1,200g for 1.5 h at room temperature. Finally, wells were washed once with PBS, activated T cells were loaded into the wells, and the plate was once again centrifuged at 1,200g for 1 h with reduced braking speed. T cells were then placed into the 37 °C incubator.

IL-2 release assay. Transduced cells were rested by culturing them for more than 10 days in the presence of 30 U ml^{-1} IL-2 added every other day for maintenance. Doxycycline (200 ng ml^{-1}) was added to the cells, which were then incubated for 6 h. Cells were washed and 5×10^4 human primary CD4$^+$ cells (transduced with the pause switch constructs or untransduced) were placed in a 96-well plate with 200 μl human growth media with activation agents added (10 ng ml^{-1} PMA + 0.5 μM ionomycin, magnetic Dynabeads coated with anti-CD3/anti-CD28 (beads:cells ratio, 0.3:1), or Raji B cells loaded with a superantigen cocktail). Doxycycline (200 ng ml^{-1}) was added into appropriate wells. After 24 h of incubation at 37 °C, the released IL-2 in the supernatant was measured with a human IL-2 ELISA kit II (BD Biosciences).

Human primary CD4$^+$ T-cell proliferation assay. Resting primary CD4$^+$ T cells were pre-treated with 200 ng ml^{-1} doxycycline for 6 h and then labelled with CellTrace Violet dye (Invitrogen). Fifty thousand dye-labelled human CD4$^+$ T cells were placed in a 96-well plate with 200 μl human growth media in the presence or absence of doxycycline (200 ng ml^{-1}) and Dynabeads coated with anti-CD3/anti-CD28 (beads:cells ratio, 0.3:1) to induce proliferation. After incubation at 37 °C for 4 days, the cells were assayed by flow cytometry. FlowJo curve fitting software was used to quantify cell proliferation as indicated by dilution of the CellTrace Violet dye in proliferating cells.

31. Peisajovich, S. G., Garbarino, J. E., Wei, P. & Lim, W. A. Rapid diversification of cell signaling phenotypes by modular domain recombination. *Science* **328,** 368–372 (2010).

32. Lee, P. J., Helman, N. C., Lim, W. A. & Hung, P. J. A microfluidic system for dynamic yeast cell imaging. *Biotechniques* **44,** 91–95 (2008).

LETTER

NLRP6 negatively regulates innate immunity and host defence against bacterial pathogens

Paras K. Anand[1], R. K. Subbarao Malireddi[1], John R. Lukens[1], Peter Vogel[2], John Bertin[3], Mohamed Lamkanfi[4,5] & Thirumala-Devi Kanneganti[1]

Members of the intracellular nucleotide-binding and oligomerization domain (NOD)-like receptor (NLR) family contribute to immune responses through activation of nuclear factor-κB (NF-κB), type I interferon and inflammasome signalling[1]. Mice lacking the NLR family member NLRP6 were recently shown to be susceptible to colitis and colorectal tumorigenesis[2–4], but the role of NLRP6 in microbial infections and the nature of the inflammatory signalling pathways regulated by NLRP6 remain unclear. Here we show that *Nlrp6*-deficient mice are highly resistant to infection with the bacterial pathogens *Listeria monocytogenes*, *Salmonella typhimurium* and *Escherichia coli*. Infected *Nlrp6*-deficient mice had increased numbers of monocytes and neutrophils in circulation, and NLRP6 signalling in both haematopoietic and radioresistant cells contributed to increased susceptibility. *Nlrp6* deficiency enhanced activation of mitogen-activated protein kinase (MAPK) and the canonical NF-κB pathway after Toll-like receptor ligation, but not cytosolic NOD1/2 ligation, *in vitro*. Consequently, infected *Nlrp6*-deficient cells produced increased levels of NF-κB- and MAPK-dependent cytokines and chemokines. Thus, our results reveal NLRP6 as a negative regulator of inflammatory signalling, and demonstrate a role for this NLR in impeding clearance of both Gram-positive and -negative bacterial pathogens.

Despite the availability of antibiotics, bacterial infections continue to threaten public health worldwide. According to estimates from the US Centers for Disease Control and Prevention, roughly 48 million people in the United States get ill every year owing to foodborne illnesses, resulting in 3,000 deaths annually[5]. *Listeria* and *Salmonella* belong to the top three pathogens contributing to foodborne infections resulting in death[5,6]. *L. monocytogenes* and *S. typhimurium* are facultative intracellular pathogens that can survive and replicate within macrophages and dendritic cells. *Listeria* rapidly escapes the phagosome to replicate in the cytosol, whereas *Salmonella* lacks a phagosomal escape mechanism and inhabits the *Salmonella*-containing vacuole[7–9]. A deeper understanding of the mechanisms by which these and other bacterial pathogens are sensed by the immune system may contribute to new approaches for developing antimicrobials. Cells of the innate immune system detect microorganisms by means of a limited set of evolutionary conserved pattern recognition receptors (PRRs)[10]. NLR proteins represent a family of intracellular PRRs that survey the cytoplasmic compartment for infectious agents and cellular damage[1]. NLRs such as NOD1 and NOD2 contribute to host defence against microbial pathogens by inducing the production of pro-inflammatory cytokines through activation of NF-κB and MAPK signalling, whereas NLRs such as NLRP1, NLRP3 and NLRC4 facilitate the activation of inflammatory caspases in large multiprotein complexes termed inflammasomes[11,12]. Our knowledge of the former NLRs has markedly improved in recent years, but initial characterization of NLRP6 has only recently been reported[2–4,13]. These studies demonstrated that NLRP6 contributes to protection against colitis, colorectal tumorigenesis and non-alcoholic steatosis by regulating the integrity of the epithelial barrier and by altering the composition of the gut microflora[2–4,13]. However, the role of NLRP6 in host defence against microbial pathogens, and the signalling pathways it regulates remain unclear.

To characterize NLRP6 expression in immune cells, *Nlrp6* transcript abundance in immune and epithelial cells was analysed by real-time quantitative PCR (qPCR). As previously reported[14], neutrophils and T cells showed the highest expression levels for *Nlrp6*, followed by macrophages, epithelial cells and dendritic cells (Supplementary Fig. 1), suggesting that NLRP6 may have an important role in these cell types. The domain architecture of NLRP6 resembles that of NLRP3, and consists of an amino-terminal Pyrin domain, a centrally located nucleotide-binding domain and a stretch of carboxy-terminally located leucine-rich repeats[1]. *Nlrp6*-deficient mice have been described[2–4,13], and animals appeared healthy and did not display gross abnormalities when housed in a specific-pathogen-free facility (data not shown). To characterize the role of NLRP6 during microbial infections, wild-type (C57BL/6J) and *Nlrp6*$^{-/-}$ mice were infected intraperitoneally (i.p.) with a lethal dose of 10^6 colony-forming units (c.f.u.) of *L. monocytogenes*. Whereas the entire cohort of wild-type mice succumbed to infection within 5–6 days, 75% of *Nlrp6*$^{-/-}$ mice survived the infection and were still alive at day 20 after infection (Fig. 1a). In agreement, wild-type mice lost on average 20% of their initial body weight before succumbing to infection, whereas in *Nlrp6*$^{-/-}$ mice weight loss was initially limited (<10% by day 3) and body weight gradually returned to normal levels after day 4 (Fig. 1b). To examine whether differential mortality in wild-type and *Nlrp6*$^{-/-}$ mice was associated with differences in bacterial dissemination, bacterial burdens in systemic organs were determined at days 1 and 3 after infection. Notably, both the liver and spleen of *Nlrp6*$^{-/-}$ mice contained significantly less bacteria than those of wild-type mice at day 1 (Fig. 1c, d); this difference increased further by day 3 after infection (Fig. 1e, f). Immunohistochemical analysis of liver sections confirmed *L. monocytogenes* to be nearly absent from the liver capsules of *Nlrp6*$^{-/-}$ mice, whereas those of wild-type mice were morphologically distorted owing to high pathogen counts (Fig. 1g). *L. monocytogenes* infection is known to trigger histopathological lesions and the formation of inflammatory cell foci, the size of which correlates with disease severity[15]. To examine the extent of immune cell infiltration, haematoxylin and eosin (H&E)-stained liver sections of *L. monocytogenes*-infected wild-type and *Nlrp6*$^{-/-}$ mice were compared. In agreement with our previous results, inflammatory cell foci in *Nlrp6*$^{-/-}$ livers were significantly smaller than those found in wild-type mice at day 3 after infection (Fig. 1h and Supplementary Fig. 2). To examine the role of NLRP6 during alternative infection routes, cohorts of wild-type and *Nlrp6*$^{-/-}$ mice were infected with *L. monocytogenes* intravenously (i.v.). As with i.p.-infected animals (Fig. 1e, f), bacterial counts in the liver and spleen of i.v.-infected *Nlrp6*$^{-/-}$ mice were significantly lower than

[1]Department of Immunology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [2]Animal Resources Center and the Veterinary Pathology Core, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [3]Pattern Recognition Receptor DPU, Immuno-Inflammation TA, GlaxoSmithKline, Collegeville, Pennsylvania 19426, USA. [4]Department of Biochemistry, Ghent University, B-9000 Ghent, Belgium. [5]Department of Medical Protein Research, VIB, B-9000 Ghent, Belgium.
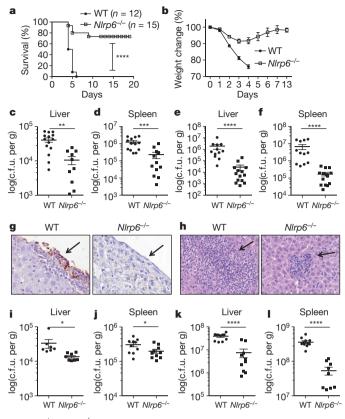
Figure 1 | *Nlrp6*−/− mice are resistant to *Listeria* and *Salmonella* infection.
a, Wild-type (WT) and *Nlrp6*−/− mice were infected i.p. with *L. monocytogenes* and survival was monitored daily for 20 days. **b**, Proportion of weight loss.
c, **d**, Bacterial loads in the liver (**c**) and spleen (**d**) on day 1. **e**, **f**, Bacterial loads in the liver (**e**) and spleen (**f**) on day 3. **g**, Livers were collected on day 3 and immunohistochemistry was performed for *L. monocytogenes* using an anti-*Listeria* antibody. Arrows indicate the extent of damage in the liver capsule of wild-type and *Nlrp6*−/− mice. **h**, Livers from *L. monocytogenes*-infected mice showing inflammatory lesions indicated by arrows. Original magnifications for all images, ×40. **i**, **j**, Wild-type and *Nlrp6*−/− mice were infected i.p. with *S. typhimurium*. Bacterial loads were determined in the liver (**i**) and spleen (**j**) on day 1. **k**, **l**, Bacterial loads in the liver (**k**) and spleen (**l**) on day 3. Each point represents an individual mouse and the mean ± s.e.m. are shown. Results show cumulative data from two different experiments. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

the burdens measured in wild-type mice at day 3 after infection (Supplementary Fig. 3). *Nlrp6*−/− mice were recently shown to have an altered microflora composition that is transferable to co-housed wild-type mice[3]. In agreement, 16S ribosomal RNA analysis confirmed an increased abundance of Bacteroidetes (Prevotellaceae family) in the gastrointestinal tract of *Nlrp6*−/− mice relative to the microbiota composition of separately housed wild-type mice (Supplementary Fig. 4a). Co-housing wild-type and *Nlrp6*−/− mice for 4 weeks equalized the prevalence of Prevotellaceae in these mice (Supplementary Fig. 4a, b). However, co-housing did not alter the resistant phenotype of *Nlrp6*−/− mice to *L. monocytogenes* infection because *Nlrp6*−/− mice contained fewer bacteria in the spleen and liver compared with the levels of co-housed wild-type mice (Supplementary Fig. 4c, d). These results suggest that NLRP6 regulates *L. monocytogenes* infection independently of its microflora composition.

To determine whether resistance of *Nlrp6*−/− mice to infection is specific to *L. monocytogenes* or extends to other pathogens, cohorts of wild-type and *Nlrp6*−/− mice were infected i.p. with *S. typhimurium*. Resistance to *S. typhimurium* infection was increased in *Nlrp6*−/− mice as illustrated by the significantly lower bacterial burdens detected in the liver and spleen at days 1 (Fig. 1i, j) and 3 (Fig. 1k, l) after infection. Inbred mice may carry distinct alleles of mouse *Nramp1* (also known

as *Slc11a1*, solute carrier family 11 and natural resistance-associated macrophage protein one), the gene products of which may differentially regulate host resistance to *Salmonella* and other bacterial pathogens[16]. *Nramp1* transcripts of both *Nlrp6*−/− and wild-type mice encoded the NRAMP1(Gly169Asp) polymorphism (data not shown), ruling out a possible effect on the increased resistance of *Nlrp6*−/− mice to *Salmonella* infection. Because both *S. typhimurium* and *L. monocytogenes* are facultative intracellular pathogens, we determined whether NLRP6 modulated resistance against extracellular pathogens such as *E. coli* as well. To this end, we analysed c.f.u. in systemic organs of wild-type and *Nlrp6*−/− mice 2 days after infection with this Gram-negative extracellular pathogen. Notably, bacterial burdens in the liver and spleen of *Nlrp6*−/− mice were significantly lower than in wild-type mice (Supplementary Fig. 5). Together, these results suggest that NLRP6 may promote systemic dissemination and growth of both Gram-negative and -positive bacteria (*S. typhimurium* and *L. monocytogenes*, respectively) as well as intracellular and extracellular bacteria (*S. typhimurium* and *E. coli*, respectively) in infected hosts.

Myeloid cells have important roles in early host defence responses against microbial pathogens. They develop in the bone marrow, but egress into the circulation after infection[17]. Because *Nlrp6* is expressed in monocytes and neutrophils (Supplementary Fig. 1), we initially made use of an automated haematology analyser—which determines blood cell types based on morphology—to characterize the influx of monocytes and neutrophils in peripheral blood of wild-type and *Nlrp6*−/− mice. Before infection, monocyte and neutrophil precursors in the bone marrow were similar (Supplementary Fig. 6), and no significant differences in the counts of circulating monocytes and neutrophils were noted in naive wild-type and *Nlrp6*−/− mice (Supplementary Fig. 7), suggesting that basal differentiation of myeloid progenitor cells was not affected in these mice. However, both the fraction and total cell counts of circulating monocytes and neutrophils were markedly higher in the peripheral blood of *L. monocytogenes*-infected *Nlrp6*−/− mice (Fig. 2a, b). To confirm these results, immune cell infiltration in the peritoneal cavity of *L. monocytogenes*-infected animals was analysed by flow cytometry. Basal monocyte and neutrophil populations in the peritoneal cavity of naive wild-type and *Nlrp6*−/− mice did not differ (data not shown). By contrast, significantly higher infiltration of monocytes and neutrophils was observed in the peritoneal cavity of *L. monocytogenes*-infected *Nlrp6*−/− mice compared with wild-type mice (Fig. 2c, d and Supplementary Fig. 8a, b). Unlike granulocyte populations, the numbers of circulating lymphocytes and infiltrated T cells in the peritoneal cavity were similar in *L. monocytogenes*-infected wild-type and *Nlrp6*−/− mice (Supplementary Fig. 8c, d). Concurrently, *Nlrp6* deficiency did not alter the induction of T-cell responses because the frequency and number of ovalbumin (OVA)-specific CD8+ T cells were similar in wild-type and *Nlrp6*−/− mice that were infected with OVA-expressing *Listeria* (Supplementary Fig. 9). Taken together, these results suggest that NLRP6 inhibits the influx of monocytes and neutrophils to the circulation and to the peritoneum during bacterial infections.

To examine further the cell types responsible for NLRP6-mediated dissemination of bacterial infections, we created wild-type and *Nlrp6*−/− bone marrow chimaeras. Bone marrow reconstitution was confirmed to reach a level of approximately 95% in irradiated mice (data not shown). As expected, *Nlrp6*−/− mice that were transplanted with *Nlrp6*−/− bone marrow were more resistant to *L. monocytogenes* infection than wild-type mice receiving wild-type bone marrow (Fig. 2e–g). Notably, the wild-type mice transplanted with *Nlrp6*−/− bone marrow and the *Nlrp6*−/− mice that received wild-type bone marrow showed an intermediate level of protection against *L. monocytogenes* (Fig. 2e–g). Indeed, weight loss (Fig. 2e) and bacterial dissemination to the spleen and liver (Fig. 2f, g) of these two chimaera groups were in between the levels seen for the two control groups (Fig. 2e–g). These results suggest that both haematopoietic and
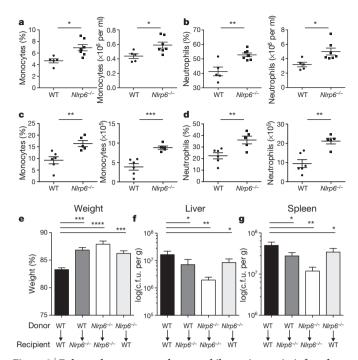
Figure 2 | Enhanced monocyte and neutrophil recruitment in infected
Nlrp6[−/−] **mice. a, b,** Wild-type and *Nlrp6*[−/−] mice were infected i.p. with *L. monocytogenes*. Blood was analysed for different cell populations based on their morphology. **c, d,** At 6 h after infection, peritoneal lavage was performed on mice infected as above, and the cells collected were analysed as CD11b[+] Gr1[−] (also known as Itgam[+] Ly6g[−]; monocytes) and CD11b[+] Gr1[+] (neutrophils) cells. **e,** Bone marrow chimaera mice were generated for *Nlrp6* as described in Methods. The percentage of weight loss in four groups of chimaeric mice is shown. **f, g,** Different groups of chimaeric mice were infected with *L. monocytogenes* and bacterial loads were determined in the liver (**f**) and spleen (**g**) on day 3. Data show mean ± s.e.m. of a representative experiment. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

non-haematopoietic cells contribute to NLRP6-mediated inhibition of bacterial clearance.

To gain understanding of the signalling mechanisms by which NLRP6 regulates bacterial infections, we first examined bacterial clearance by *in vitro* cultured *Nlrp6*[−/−] macrophages. Phagocytosis and bacterial replication of *L. monocytogenes* in *Nlrp6*[−/−] bone marrow-derived macrophages (BMDMs) were comparable to those of wild-type macrophages (Supplementary Fig. 10). Unlike results from overexpression studies[14], but in agreement with a recent report that used *NLRP6* short interfering RNA (siRNA) knockdown in human PBMCs[18], we found NLRP6 to be dispensable for inflammasome activation by *L. monocytogenes* and *S. typhimurium* because neither caspase-1 processing nor maturation of IL-1β was affected in *Nlrp6*-deficient macrophages infected with these pathogens or stimulated with the bacterial ligands lipopolysaccharide (LPS) and the synthetic lipoprotein Pam3CSK4 in the presence of ATP (Supplementary Fig. 11).

We next assessed the potential role of NLRP6 in regulating activation of NF-κB and MAPK signalling in macrophages infected with *L. monocytogenes*. In wild-type BMDMs, the phosphorylation of extracellular signal-regulated kinases 1 and 2 (ERK1/2) and IκB peaked 60 min after infection and gradually diminished thereafter (Fig. 3a, b). Infected *Nlrp6*[−/−] macrophages had significantly increased phospho-ERK1/2 levels, and failed to downregulate phospho-IκB (Fig. 3a, b). In addition to macrophages, *Nlrp6*[−/−] neutrophils also responded to *L. monocytogenes* infection with increased phospho-IκB and phospho-ERK1/2 levels (Supplementary Fig. 12a), suggesting that NLRP6 may negatively regulate NF-κB and ERK activation downstream of TLRs. In agreement, the TLR2 ligand Pam3CSK4 also induced increased phospho-ERK1/2 and phospho-IκB levels in *Nlrp6*[−/−] BMDMs (Fig. 3c, d). Similar observations were made using
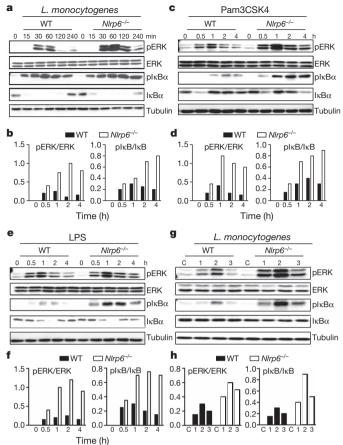

Figure 3 | NLRP6 negatively regulates NF-κB and ERK signalling.
a, c, e, Wild-type and *Nlrp6*[−/−] macrophages were exposed to *L. monocytogenes* or to Pam3CSK4 or LPS, and samples were immunoblotted with the indicated antibodies. **b, d, f,** Bar graphs showing semi-quantification of phosphorylated ERK (pERK) and pIκB levels relative to ERK and IκB levels, respectively. **g,** Liver lysates from wild-type and *Nlrp6*[−/−] mice were examined for activation of NF-κB and MAPK signalling by western blot analysis. Each lane represents an individual mouse; C denotes an uninfected mouse. **h,** Bar graphs showing semi-quantification of pERK and pIκB levels relative to ERK and IκB levels, respectively.

the TLR4 ligand LPS (Fig. 3e, f). Increased phospho-IkB and phospho-ERK1/2 levels were also noted in liver lysates of infected *Nlrp6*[−/−] mice (Fig. 3g, h). Unlike with TLR agonists, phospho-IκB and phospho-ERK1/2 levels induced after ligation of the cytosolic receptors NOD1 and NOD2 were similar in wild-type and *Nlrp6*[−/−] cells (Supplementary Fig. 12b, c), suggesting that NLRP6 is specifically involved in negatively regulating TLR-induced NF-κB and MAPK activation. Thus, in marked contrast to overexpression studies[14], our results in *Nlrp6*[−/−] cells suggest that NLRP6 may diminish rather than enhance TLR-induced activation of NF-κB and ERK signalling. To confirm the role of NLRP6 in regulating these pathways further, we measured the production of NF-κB- and MAPK-dependent cytokine and chemokine levels at different time points. Abundance of tumour necrosis factor (TNF)-α, interleukin (IL)-6 and keratinocyte chemoattractant (KC) transcripts in cell lysates and secreted cytokines in culture supernatants were significantly higher in *Nlrp6*[−/−] macrophages infected with *L. monocytogenes*, or stimulated with Pam3CSK4 and LPS (Supplementary Fig. 13a–d). Interferon-β transcript levels nearly doubled in LPS-stimulated *Nlrp6*[−/−] macrophages after 2 h stimulation, but subsequently returned to levels of wild-type cells (Supplementary Fig. 14). Cytotoxicity was limited to baseline levels in both wild-type and *Nlrp6*[−/−] macrophages, ruling out an important role for cell death induction in differential secretion of cytokines and chemokines (Supplementary Fig. 15). Increased levels

of IL-6 and KC were also measured in serum and the peritoneal cavity of infected $Nlrp6^{-/-}$ mice (Supplementary Fig. 13e, f). These results suggest that NLRP6 specifically suppresses TLR-induced NF-κB and MAPK activation during *L. monocytogenes* infection *in vitro* and *in vivo*.

NF-κB belongs to a family of nuclear transcriptional regulators that initiate pro-inflammatory gene expression[19]. To determine whether canonical or non-canonical NF-κB activation is altered in $Nlrp6$-deficient cells, we analysed the phosphorylation status of the non-canonical NF-κB effector p100, whereas p105 phosphorylation status was monitored as a parameter of canonical NF-κB signalling. *L. monocytogenes* infection induced significantly higher levels of phosphorylated p105 in $Nlrp6^{-/-}$ macrophages, whereas those of phosphorylated p100 were similar in wild-type and $Nlrp6$-deficient macrophages (Fig. 4a). Similar results were obtained when analysing the isolated cytosolic fraction of *L. monocytogenes*-infected cells (Fig. 4a). Furthermore, this correlated with increased translocation
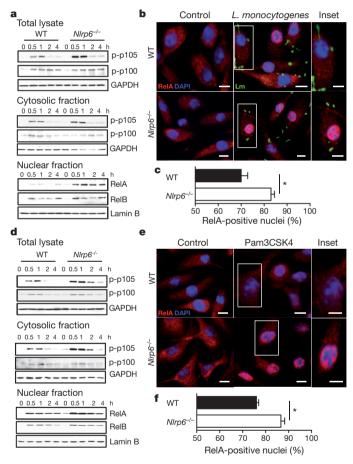
of the NF-κB subunit RelA (also known as p65) in the nuclear fraction of infected $Nlrp6$-deficient cells, whereas the levels of the non-canonical effector RelB were similar in wild-type and $Nlrp6^{-/-}$ macrophages (Fig. 4a). Confocal immunofluorescence microscopy of *L. monocytogenes*-infected wild-type and $Nlrp6^{-/-}$ cells confirmed the increased nuclear translocation of RelA in the latter genotype (Fig. 4b, c). Increased p105 phosphorylation was also evident in total and cytosolic lysates of Pam3CSK4-stimulated $Nlrp6^{-/-}$ macrophages (Fig. 4d). In addition, enhanced RelA translocation was confirmed by western blotting in nuclear lysates and by confocal immunofluorescence staining (Fig. 4d–f). qPCR analysis of interleukin-1 receptor-associated kinase M (IRAK-M) and the NF-κB negative regulator A20 failed to demonstrate differential expression of these negative regulators of TLR signalling in LPS- and Pam3CSK4-stimulated $Nlrp6$-deficient cells (Supplementary Fig. 16). Although further analysis is required to determine the precise level at which NLRP6 negatively regulates TLR-induced responses, our results indicate that NLRP6 suppresses TLR-induced MAPK and canonical NF-κB signalling to dampen the production of pro-inflammatory cytokines and chemokines during bacterial infection. NLRP6 activation consequently leads to increased susceptibility to both intracellular (*L. monocytogenes* and *S. typhimurium*) and extracellular (*E. coli*) as well as Gram-positive (*L. monocytogenes*) and -negative (*E. coli* and *S. typhimurium*) bacterial pathogens. Our bone marrow chimaera experiments suggested that the haematopoietic and non-haematopoietic compartments contribute to the increased resistance of $Nlrp6^{-/-}$ mice to infection. We further demonstrated here that NLRP6 has a crucial role in regulating host and immune responses in cells of the myeloid lineage, and further analysis is required to determine the specific non-haematopoietic cell type(s) contributing to NLRP6-mediated host defence signalling *in vivo*.

The detrimental role of NLRP6 during systemic infection with bacterial pathogens appears in sharp contrast to its protective role in gastrointestinal tract, where it contributes to maintaining a healthy composition of the gut microbiota[20]. We propose that these apparently contradictory roles of NLRP6 may be explained by the spatiotemporal context in which it operates. As NLRP6 seems to be essential for dampening the production of pro-inflammatory cytokines and chemokines downstream of TLRs, it may have a protective role under conditions in which strong inflammatory responses might be detrimental to the host (as during colitis). On the other hand, NLRP6 may play an adverse role during situations such as bacterial infections in which potent inflammatory responses would be protective to the host.

Notably, previous studies characterizing the NLR family members NOD1, NOD2, NLRC4 and NLRP3 all pointed to a protective role for these NLRs in clearing microbial pathogens, and deficiency in these NLRs led to increased morbidity and mortality in infected animals[21–24]. By contrast, our data suggest that NLRP6 dampens inflammatory signalling, thereby promoting bacterial dissemination and colonization of systemic organs of the host. In light of these findings, we propose that one subclass of NLRs (including NOD1, NOD2 and NLRC4) may directly respond to microbial pathogens detected in cytosolic compartment by initiating pro-inflammatory signalling that eventually contributes to bacterial clearance. A second subclass of NLR proteins, such as NLRP6, may act as a molecular switch to dampen or silence TLR-induced pathways triggered by extracellular recognition of bacterial ligands to prevent overt pathology. Regardless, our results highlight a crucial role for NLRP6 in dampening host responses against several bacterial pathogens and suggest that therapeutic inhibition of NLRP6 activation may prove beneficial for treating infectious diseases.

METHODS SUMMARY

Figure 4 | NLRP6 negatively regulates canonical NF-κB activation pathway. a, BMDMs from wild-type and $Nlrp6^{-/-}$ mice were infected with *L. monocytogenes* and the phosphorylation of p105 and p100 was examined in total cell lysate and the cytosolic fraction. The nuclear fraction was examined for RelA and RelB activation in wild-type and $Nlrp6^{-/-}$ cells. **b**, Cells grown on coverslips were infected with green fluorescent protein (GFP)-labelled *L. monocytogenes* (Lm) and examined for nuclear translocation of RelA using an anti-RelA antibody (red). DAPI, 4′,6-diamidino-2-phenylindole. Scale bars, 10 μm. **c**, Quantitative analysis of RelA-positive nuclei. **d**, Wild-type and $Nlrp6^{-/-}$ macrophages were stimulated with Pam3CSK4, and the phosphorylation of p105 and p100 was examined in total cell lysate and the cytosolic fraction. The nuclear fraction was examined for RelA and RelB activation in wild-type and $Nlrp6^{-/-}$ cells. **e**, Cells grown on coverslips were stimulated with Pam3CSK4 and examined for nuclear translocation of RelA using an anti-RelA antibody (red). Scale bars, 10 μm. **f**, Quantitative analysis of RelA-positive nuclei. Data show mean ± s.e.m. from three different experiments. *$P < 0.05$.

conducted under protocols approved by the St. Jude Children's Research Hospital Committee on Use and Care of Animals. Lethal *L. monocytogenes* infection was established by infecting female mice with 1×10^6 c.f.u. bacteria administered i.p. To study bacterial clearance, mice were infected with *L. monocytogenes* at a dose of 3×10^5 c.f.u. (i.p.). Immune cell infiltration in the circulation of infected mice was performed on blood collected by submandibular venipuncture and analysed on a Forcyte haematology analyser. Bone marrow chimaera experiments were done as described previously[25]. Bone marrow macrophages were prepared and infected as described previously[7]. Samples were run on SDS–PAGE and immunoblotted with antibodies as reported previously[7]. Data are expressed as mean ± s.e.m. Differences were analysed by a Student's *t*-test. *P* values ≤ 0.05 were considered significant.

Full Methods and any associated references are available in the online version of the paper.

1. Kanneganti, T. D., Lamkanfi, M. & Nunez, G. Intracellular NOD-like receptors in host defense and disease. *Immunity* **27,** 549–559 (2007).
2. Chen, G. Y., Liu, M., Wang, F., Bertin, J. & Nunez, G. A functional role for Nlrp6 in intestinal inflammation and tumorigenesis. *J. Immunol.* **186,** 7187–7194 (2011).
3. Elinav, E. *et al.* NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* **145,** 745–757 (2011).
4. Normand, S. *et al.* Nod-like receptor pyrin domain-containing protein 6 (NLRP6) controls epithelial self-renewal and colorectal carcinogenesis upon injury. *Proc. Natl Acad. Sci. USA* **108,** 9601–9606 (2011).
5. Centers for Disease Control and Prevention. CDC Estimates of Foodbornes Illness in the United States, 2011 Estimates. http://www.cdc.gov/foodborneburden/2011-foodborne-estimates.html.
6. Scallan, E. *et al.* Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* **17,** 7–15 (2011).
7. Anand, P. K. *et al.* TLR2 and RIP2 pathways mediate autophagy of *Listeria monocytogenes* via ERK activation. *J. Biol. Chem.* **286,** 42981–42991 (2011).
8. Corr, S. C. & O'Neill, L. A. *Listeria monocytogenes* infection in the face of innate immunity. *Cell. Microbiol.* **11,** 703–709 (2009).
9. Hybiske, K. & Stephens, R. S. Exit strategies of intracellular pathogens. *Nature Rev. Microbiol.* **6,** 99–110 (2008).
10. Creagh, E. M. & O'Neill, L. A. TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends Immunol.* **27,** 352–357 (2006).
11. Kanneganti, T. D. Central roles of NLRs and inflammasomes in viral infection. *Nature Rev. Immunol.* **10,** 688–698 (2010).
12. Pétrilli, V., Dostert, C., Muruve, D. A. & Tschopp, J. The inflammasome: a danger sensing complex triggering innate immunity. *Curr. Opin. Immunol.* **19,** 615–622 (2007).
13. Henao-Mejia, J. *et al.* Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. *Nature* **482,** 179–185 (2012).
14. Grenier, J. M. *et al.* Functional screening of five PYPAF family members identifies PYPAF5 as a novel regulator of NF-κB and caspase-1. *FEBS Lett.* **530,** 73–78 (2002).
15. Stevenson, M. M., Kongshavn, P. A. & Skamene, E. Genetic linkage of resistance to *Listeria monocytogenes* with macrophage inflammatory responses. *J. Immunol.* **127,** 402–407 (1981).
16. Lara-Tejero, M. *et al.* Role of the caspase-1 inflammasome in *Salmonella typhimurium* pathogenesis. *J. Exp. Med.* **203,** 1407–1412 (2006).
17. Serbina, N. V. & Pamer, E. G. Monocyte emigration from bone marrow during bacterial infection requires signals mediated by chemokine receptor CCR2. *Nature Immunol.* **7,** 311–317 (2006).
18. Meixenberger, K. *et al. Listeria monocytogenes*-infected human peripheral blood mononuclear cells produce IL-1β, depending on listeriolysin O and NLRP3. *J. Immunol.* **184,** 922–930 (2010).
19. Waterfield, M. R., Zhang, M., Norman, L. P. & Sun, S. C. NF-κB1/p105 regulates lipopolysaccharide-stimulated MAP kinase signaling by governing the stability and function of the Tpl2 kinase. *Mol. Cell* **11,** 685–694 (2003).
20. van Ooij, C. Immunology: NLRP6 keeps the bad bacteria at bay. *Nature Rev. Microbiol.* **9,** 481 (2011).
21. Girardin, S. E. *et al.* Nod1 detects a unique muropeptide from Gram-negative bacterial peptidoglycan. *Science* **300,** 1584–1587 (2003).
22. Kobayashi, K. S. *et al.* Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* **307,** 731–734 (2005).
23. Franchi, L. *et al.* Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1β in salmonella-infected macrophages. *Nature Immunol.* **7,** 576–582 (2006).
24. Broz, P. *et al.* Redundant roles for inflammasome receptors NLRP3 and NLRC4 in host defense against *Salmonella. J. Exp. Med.* **207,** 1745–1755 (2010).
25. Zaki, M. H. *et al.* The NLRP3 inflammasome protects against loss of epithelial integrity and mortality during experimental colitis. *Immunity* **32,** 379–391 (2010).

Author Contributions T.-D.K., M.L. and P.K.A. designed the research; P.K.A., R.K.S.M. and J.R.L. performed the experiments; P.V. performed and analysed the histopathology data; J.B. provided essential reagents; T.-D.K., M.L., P.K.A., R.K.S.M., J.R.L., P.V. and J.B. analysed the data; P.K.A., M.L. and T.-D.K. wrote the paper; and T.-D.K. conceived the study, designed the experiments and provided overall direction. M.L. and T.-D.K. contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.-D.K. (thirumala-devi.kanneganti@stjude.org).

METHODS

Bacterial infections and enumeration of bacterial burdens. Lethal *L. monocytogenes* infection was established by infecting female mice with 1×10^6 c.f.u. bacteria administered i.p. Animals were weighed and monitored daily for mortality for up to 20 days. Differences in survival were analysed by the Mantel–Cox test. To study bacterial clearance, mice were infected with *L. monocytogenes*, at a dose of either 3×10^5 c.f.u. (i.p.) or 3×10^4 c.f.u. (i.v.)[7]. Alternatively, mice were infected with *S. typhimurium* at 1×10^4 c.f.u. (i.p). *E. coli* infection was performed in 6-week-old female mice with 5×10^7 c.f.u. (i.p.). Mice were euthanized at days 1 and 3 after infection (or day 2 in the case of *E. coli* infection), liver and spleen tissue were homogenized and bacterial burdens were enumerated by serial dilution on brain–heart infusion (BHI) or lysogeny broth (LB) agar plates as described previously[7].

Histopathology. Formalin-preserved liver sections were processed and embedded in paraffin by standard techniques. Longitudinal sections of 5-μm thick were stained with haematoxylin and eosin and examined by a pathologist blinded to the experimental groups. For immunohistochemistry, formalin-fixed paraffin-embedded tissues were cut into 4 μm sections and slides were stained with an anti-*Listeria* antibody.

Analysis of immune cell infiltration in the peritoneum and in circulation. Wild-type and $Nlrp6^{-/-}$ female mice were infected with 3×10^5 *L. monocytogenes* (i.p.). At 3 h after infection, blood was collected by submandibular venipuncture and analysed on a Forcyte haematology analyser (Oxford Sciences) by a haematologist blinded to the experimental groups. In addition, cells were collected by peritoneal lavage in either 1 ml IMDM (for cytokines analysis) or 10 ml IMDM (for total cell recruitment to the cavity) on ice. Cells obtained from wild-type and $Nlrp6^{-/-}$ mice were washed twice with PBS and surface-stained for anti-CD11b (Biolegend; clone M1/70), anti-Gr1 (Biolegend; clone RB6-8C5), anti-TCR-β (Biolegend; clone H57-597) and anti-Ly6G (BD Biosciences). Monocytes were characterized as CD11b$^+$ Gr1$^-$ TCR-β$^-$ cells, whereas CD11b$^+$ Gr1$^+$ cells were taken as neutrophils.

Wild-type and $Nlrp6^{-/-}$ mice were infected i.v. with 3×10^4 c.f.u. of *L. monocytogenes* that was engineered to express the OVA antigen. At the peak of the T-cell response (day 8), 100 μl of peripheral blood was obtained and peripheral blood leukocytes were isolated after red blood cell lysis. Cells were FACS-stained with OVA-tetramers and T-cell lineage markers (anti-CD4 and anti-CD8).

Bone marrow chimaeras. Bone marrow transfer was used to create $Nlrp6^{-/-}$ chimaera mice in which the genetic deficiency of $Nlrp6$ was contained to either the circulating cells ($Nlrp6^{-/-}$ > wild-type chimaera) or non-haematopoietic tissue (wild type > $Nlrp6^{-/-}$)[25]. In brief, bone marrow cells were collected from the femur and tibia of congenic wild-type (expressing CD45.1 leukocyte antigen) or $Nlrp6^{-/-}$ (expressing CD45.2 leukocyte antigen) donor mice by flushing with HBSS. After thorough washing, cells were resuspended in PBS and 100 μl of this cell suspension containing 10^7 cells was injected retro-orbitally to the irradiated host mice. Four chimaera groups were generated (wild type > wild type; wild type > $Nlrp6^{-/-}$; $Nlrp6^{-/-}$ > wild type; and $Nlrp6^{-/-}$ > $Nlrp6^{-/-}$). The use of CD45.1-expressing congenic wild-type mice facilitated the verification of bone marrow reconstitution at 6 weeks after irradiation. Cells in the blood were labelled with allophycocyanin (APC)-conjugated CD45.1 and fluorescein isothiocyanate (FITC)-conjugated CD45.2 antibody. The chimaera mice had ~95% reconstitution with the donor bone marrow.

Macrophage culture and *in vitro* stimulation. BMDMs and peritoneal macrophages were prepared as described previously[7]. Cells were subsequently infected with *L. monocytogenes* or *S. typhimurium* as indicated; or stimulated with Pam3CSK4 (10 μg ml^{-1}) or LPS (100 ng ml^{-1}). At the indicated time points, supernatant was collected and stored at $-80\,^{\circ}$C for cytokine analysis at a later time and cells were lysed in RIPA buffer supplemented with complete protease inhibitor cocktail (Roche) and PhosSTOP (Roche). For analysis of caspase-1 activation and IL-1β secretion, macrophages were cultured with bacteria or TLR ligands for 3 h followed by addition of 5 mM ATP (Sigma) for 30 min.

***In vivo* lysates and serum cytokines.** Liver was obtained from control and infected wild-type and $Nlrp6^{-/-}$ mice at 3 h after infection. The tissue was homogenized in RIPA buffer supplemented with complete protease inhibitor cocktail (Roche) and PhosSTOP (Roche). At this time, blood was also collected by submandibular venipuncture for serum cytokine analysis.

ELISA. Cytokine ELISA was performed according to manufacturer's instructions (Milliplex).

LDH assay. LDH assay was performed by using the Promega cytotoxicity kit according to manufacturer's instructions.

Western blotting. Lysates were prepared in RIPA buffer supplemented with complete protease inhibitor cocktail (Roche) and PhosSTOP (Roche). Samples were resolved by SDS–PAGE and transferred to polyvinylidene difluoride (PVDF) membranes by electroblotting. Membranes were blocked in 5% non-fat milk and incubated overnight with primary antibodies at 4 °C and subsequently with horseradish peroxidase (HRP)-tagged secondary antibodies at room temperature for 1 h. Immunoreactive proteins were visualized with ECL method (Pierce). Semi-quantitative analysis of western blot bands was performed by Image J program.

Confocal immunofluorescence microscopy. Cells were grown overnight on glass coverslips. The next day, wild-type and $Nlrp6^{-/-}$ macrophages were either infected with GFP-labelled *L. monocytogenes* or stimulated with Pam3CSK4. The cells were fixed with 4% paraformaldehyde and processed for microscopy with an anti-RelA antibody (Santa-Cruz) as described previously[7]. Cells, grown on coverslips, were mounted on glass slides with ProLong gold antifade reagent (Invitrogen) and visualized with a Nikon C1 confocal microscope using a ×40 objective lens. The images were processed and analysed with Image J program.

Real-time PCR. Total RNA was isolated with Trizol (Invitrogen) according to the manufacturer's instructions. For $Nlrp6$ expression, macrophages and dendritic cells were differentiated from bone marrow cells, neutrophils were obtained from the peritoneal cavity after 5 h of thioglycollate injection, T cells were purified from the spleen and epithelial cells were isolated from the intestine. One microgram of RNA was reverse-transcribed to complementary DNA with poly dT primers using the first-strand cDNA synthesis kit (Invitrogen). Transcript levels of different genes were analysed using SYBR-green (Applied Biosystems) on an ABI7500 real-time PCR instrument. Expression of *Gapdh* was used as an internal control. Messenger RNA expression levels were normalized against a standard curve.

Co-housing experiments and analysis of fecal flora. Age and gender-matched wild-type and $Nlrp6^{-/-}$ mice were co-housed together at 1:1 ratio for 4 weeks as described previously[3]. Fecal samples were obtained from the two groups of mice before and after co-housing. Frozen fecal samples were processed for DNA isolation using QiaAmp stool kit (Qiagen). Ten nanograms of DNA was used in the PCR reaction and previously reported primer pairs were used for analysis of 16S rRNA analysis[3].

Intracellular growth curve. Wild-type and $Nlrp6^{-/-}$ macrophages were infected with *L. monocytogenes* with a multiplicity of infection of 1:1 and intracellular growth curves were generated as described previously[26].

Statistics. Differences in the survival of animals were calculated by log-rank Mantel–Cox test using Graphpad Prism software. In all other experiments, *P* values were calculated using the two-tailed Student's *t*-test. *P* values are denoted by *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$.

26. Portnoy, D. A., Jacks, P. S. & Hinrichs, D. J. Role of hemolysin for the intracellular growth of *Listeria monocytogenes*. *J. Exp. Med.* **167,** 1459–1471 (1988).

LETTER

APJ acts as a dual receptor in cardiac hypertrophy

Maria Cecilia Scimia[1]*, Cecilia Hurtado[1]*, Saugata Ray[1], Scott Metzler[2], Ke Wei[1], Jianming Wang[1], Chris E. Woods[3], Nicole H. Purcell[4], Daniele Catalucci[5,6], Takeshi Akasaka[1], Orlando F. Bueno[7], George P. Vlasuk[7]†, Perla Kaliman[8], Rolf Bodmer[1], Layton H. Smith[1], Euan Ashley[3], Mark Mercola[1], Joan Heller Brown[4] & Pilar Ruiz-Lozano[1,2]

Cardiac hypertrophy is initiated as an adaptive response to sustained overload but progresses pathologically as heart failure ensues[1]. Here we report that genetic loss of APJ, a G-protein-coupled receptor, confers resistance to chronic pressure overload by markedly reducing myocardial hypertrophy and heart failure. In contrast, mice lacking apelin (the endogenous APJ ligand) remain sensitive, suggesting an apelin-independent function of APJ. Freshly isolated APJ-null cardiomyocytes exhibit an attenuated response to stretch, indicating that APJ is a mechanosensor. Activation of APJ by stretch increases cardiomyocyte cell size and induces molecular markers of hypertrophy. Whereas apelin stimulates APJ to activate $G\alpha_i$ and elicits a protective response, stretch signals in an APJ-dependent, G-protein-independent fashion to induce hypertrophy. Stretch-mediated hypertrophy is prevented by knockdown of β-arrestins or by pharmacological doses of apelin acting through $G\alpha_i$. Taken together, our data indicate that APJ is a bifunctional receptor for both mechanical stretch and the endogenous peptide apelin. By sensing the balance between these stimuli, APJ occupies a pivotal point linking sustained overload to cardiomyocyte hypertrophy.

G-protein-coupled receptors (GPCRs) have been widely implicated in the control of cardiac function. These receptors couple to hetero-trimeric GTP-binding proteins of the $G\alpha_s$, $G\alpha_i$, $G\alpha_{q/11}$ and $G\alpha_{12/13}$ families, and transduce the GPCR signal to intracellular targets. Numerous studies have linked $G\alpha_s$ to increased contractility, $G\alpha_{q/11}$ to pathological hypertrophy[2,3] and $G\alpha_i$ to cardioprotection[4]. APJ is a GPCR identified as the receptor for the adipokine apelin[5,6]. Apelin-activated APJ signals through $G\alpha_i$, exerting a positive effect on cardiac contractility[7–9] and a vasodilator activity that counteracts angiotensin-II-induced atheroma[10,11]. Apelin administration blunts progression to hypertrophy (Supplementary Fig. 1 and Supplementary Tables 2 and 3) and apelin-KO mice show susceptibility to heart failure[12] (also see Supplementary Fig. 1 and Supplementary Table 1). Thus, apelin and its receptor APJ are emerging as potential therapeutic targets.

We examined the response of APJ knockout (APJ-KO) mice to sustained pressure overload by transaortic constriction (TAC). Although deletion of APJ resulted in some prenatal lethality[13,14], all viable APJ-KO mice showed normal adult appearance and cardio-vascular parameters at baseline (Supplementary Table 4). However, APJ-null animals were resistant to the pathological hypertrophic response to TAC (Fig. 1a–d) observed both in wild type (WT) and in apelin-KO mice (Supplementary Fig. 1g–i). APJ-KO mice responded to TAC by initially increasing cardiac mass, but the maladaptive progression to dilated ventricular hypertrophy was blunted shortly after injury (Supplementary Table 4). The protective effect persisted long-term (Fig. 1a, b, g, h) in all parameters measured, including diminished cardiomyocyte size (Fig. 1c, d), reduced fibrosis (Fig. 1e, f), sustained cardiac contractility (Fig. 1g) relative to WT and apelin-KO mice (Supplementary Tables 1 and 4) and reduced heart weight/body weight ratio (Fig. 1h). Baseline cardiac contractility, measured as

percentage fractional shortening, was approximately 38% across genotypes. After 90 days of TAC, percentage fractional shortening decreased to $22 \pm 2\%$ in WT, $23 \pm 1\%$ in apelin KO mice, but remained at $34 \pm 2\%$ in APJ-KO mice ($P = 0.01$ between APJ-KO
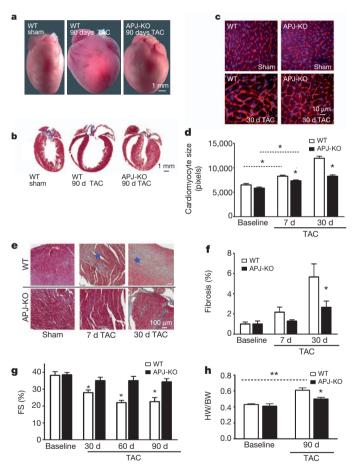

Figure 1 | APJ-KO mice are protected from hypertrophy after TAC.
a, b, Anatomical view (**a**) and histological sections (**b**) of WT and APJ-KO mice 90 days after surgery. **c,** Cell membrane staining (wheat germ agglutinin). **d,** Quantification from **c. e,** Trichrome staining (fibrosis in blue, stars). **f,** Quantification of **e. g,** Fractional shortening (FS%) decreased in WT mice after TAC, but did not change significantly in the APJ-KO mice. APJ-KO mice fail to develop heart failure upon sustained TAC as shown by echocardiographyc analysis. **h,** Heart weight/body weight ratio (HW/BW) at baseline and in TAC-operated mice, 90 days after surgery (see Supplementary Table 4 for details). Error bars, s.e.m. *$P < 0.05$ between indicated groups, analysis of variance.

[1]Sanford-Burnham Medical Research Institute, La Jolla, California 92037, USA. [2]Department of Pediatrics, School of Medicine, Stanford University, California 94305, USA. [3]Department of Medicine, School of Medicine, Stanford University, California, 94305, USA. [4]Department of Pharmacology, University of California, San Diego, California, USA. [5]Biomedical and Genetic Research Institute, National Research Council, via Fantoli 16/15, 20138, Milan, Italy. [6]Humanitas Clinical and Research Center, Via Manzoni 56, 20089 Rozzano, Milan, Italy. [7]Wyeth Pharmaceutical, Madison, New Jersey, USA. [8]Institute of Biomedical Research, August Pi i Sunyer (IDIBAPS), E-08036 Barcelona, Spain. †Present address: Sirtris, a GSK Company, Cambridge, Massachusetts 02139, USA.
*These authors contributed equally to this work.

and WT) (Fig. 1g and Supplementary Tables 1 and 4). In summary, both WT and apelin-KO mice presented clear signs of heart failure after 90 days of TAC, whereas APJ-KO mice were almost unaffected. The maintenance of cardiac function in the APJ-KO demonstrates that the expression of APJ is necessary to elicit heart failure in response to pressure overload.

The different responses of apelin-KO and APJ-KO mice to TAC imply that either apelin can act independently of APJ, or that APJ transduces a signal independently of apelin. We tested the first hypothesis by infusing APJ-KO mice with apelin (285 µg kg^{-1} per 24 h) and examining two readouts: contractility under TAC and vascular tone. Notably, apelin infusion did not increase cardiac contractility (percentage fractional shortening) in TAC-APJ-KO mice, in contrast to the characteristic improvement seen in TAC-WT animals (Supplementary Fig. 2a). In the absence of apelin infusion, endogenous levels of apelin in blood increased after TAC from 1 to 2 ng ml^{-1} and that rise was not different in WT and APJ-KO mice, making it unlikely that the protection achieved in the APJ-KO is due to hyper-activation of apelin signalling (Supplementary Fig. 2b). To test vascular tone, systolic and diastolic blood pressures were increased by infusion of Ang-II (1,000 ng kg^{-1} min^{-1}). Apelin infusion significantly decreased systolic blood pressure in WT animals but not in APJ-KO mice (Supplementary Fig. 2c–f), further indicating that apelin activity requires APJ.

Because the mechanical properties of the heart change markedly during pressure overload[15], and the structurally related angiotensin receptor (AT-1) can act as a mechanosensor[16], we asked whether APJ responds to mechanical stretch. Initially these experiments were challenging as cultured cardiomyocytes consistently downregulate the expression of endogenous APJ (Supplementary Fig. 3a), and studies had to be restricted to freshly isolated adult cardiomyocytes. We mimicked the effect of pressure overload by using a carbon fibre technique[17] to stretch cardiomyocytes and evaluated their Frank–Starling gain (FSG). FSG is a dimensionless metric of the force that can be recruited by stretch[18]. Freshly isolated adult cardiomyocytes from

WT mice showed a significantly higher FSG than cardiomyocytes from APJ-KO mice (Fig. 2a–c). Treatment with apelin decreased the FSG in WT cardiomyocytes but showed no effect in APJ-KO cells (Fig. 2d). Therefore, apelin modulated the response to stretch only in cardiomyocytes with APJ receptors.

Engineered cells stably expressing human APJ (APJ-HEK) responded to apelin by increasing the content of phosphorylated extracellular signal-regulated kinase (pERK) (Fig. 3a, b) whereas parental (HEK) cells showed no significant change. Stretch also increased pERK content in cells expressing APJ (Fig. 3a, b). pERK levels, therefore, reflect the cellular response to APJ activation by either stretch or apelin. pERK was therefore used as a simple readout of APJ activation. The Gα$_i$ inhibitor PTX blocked the ability of apelin, but not stretch, to induce APJ-dependent phosphorylation of ERK (Fig. 3a, b), first suggesting that separate mechanisms link stretch and apelin to APJ intracellular signalling.
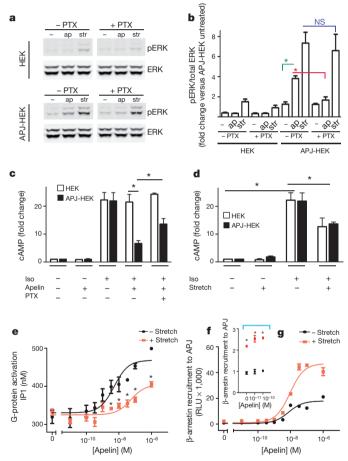
Figure 3 | **Stretch activation of APJ enhances β-arrestin while reducing G-protein signalling. a, b,** Immunoblot (**a**) and quantification (**b**) of ERK from APJ stably transfected (APJ-HEK) and parental (HEK) cells treated for 5 min with 100 nM apelin or stretch in absence or presence of PTX (Gα$_i$ inhibitor), $n = 3$. **c, d,** Effect of 1 µM apelin (**c**) or stretch (**d**) on cAMP levels. Isoproterenol (iso, 1 µM) was used to elevate cAMP artificially to study Gα$_i$ activation, $n = 4$. **e,** G-protein (Gα$_{q,s,i}$ and Gα$_{12/13}$) activation by apelin and effect of stretch (red) in CHO cells expressing APJ and Gα$_{16}$. Receptor stimulation activated the promiscuous Gα$_{16}$, phospholipase-Cβ and caused the accumulation of IP1 (representative of three experiments, $n = 4$ samples). **f, g,** Arrestin recruitment to the APJ receptor in response to apelin (black) and stretch (red) by an enzyme complementation assay in CHO cells expressing recombinant APJ and β-arrestin2 (representative of three experiments, $n = 3$ samples). **f,** Full range of β-arrestin binding to APJ under either physiological or pharmacological doses of apelin. **g,** Data points for 0–10^{-10} M apelin. RLU, relative light units. Error bars, s.e.m. *$P < 0.05$ between indicated groups, analysis of variance.
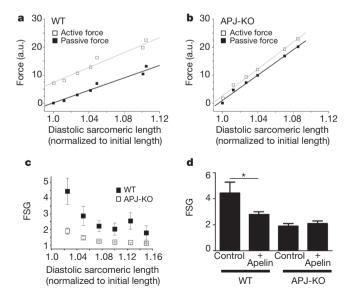

Figure 2 | **APJ mediates a stretch response that can be modulated by apelin. a, b,** Representative force measurements (arbitrary units) for end-diastolic and end-systolic length–tension relationships in adult cardiomyocytes from WT (**a**) and APJ-KO (**b**) mice plotted against diastolic length (normalized to unstretched length). Cells were paced at 1 Hz. **c,** FSG attained by dividing the active force by the passive force from experiments in **a** and **b** plotted as a function of diastolic length ($n = 8$ WT and 7 APJ-KO cardiomyocytes). **d,** Average FSG (at 1.02 sarcomeric length) from **c** is shown for APJ-KO and WT both with (+) and without (−) 10 nM apelin administration ($n = 6$ for WT + apelin and $n = 7$ for APJ-KO + apelin). Error bars, s.e.m.

We next examined whether activation of APJ by stretch or apelin differentially modulates G-protein-generated second messengers. Activation of $G\alpha_i$ is associated with inhibition of adenyl-cyclase and reduced cAMP, whereas activation of $G\alpha_s$ increases cAMP[19]. cAMP did not increase after application of stretch or apelin, arguing against activation of adenyl-cyclase (and $G\alpha_s$) (Fig. 3c, d, no isoproterenol conditions). In contrast, when isoproterenol was used to elevate the intracellular concentration of cAMP[20], apelin addition decreased cAMP levels in the APJ-HEK cells (Fig. 3c). This effect of apelin was partly inhibited by PTX, consistent with the involvement of $G\alpha_i$

(Fig. 3c). Apelin did not modify isoproterenol-stimulated cAMP formation in untransfected HEK controls, showing that the decrease is mediated by APJ. These data agree with previous reports[21] and demonstrate that $G\alpha_i$ transduces the signal initiated by apelin binding to APJ. In contrast, stretch reduced cAMP in parental HEK as well as in APJ-HEK cells (Fig. 3d), indicating that, although stretch can activate $G\alpha_i$ signalling, this response is not mediated through APJ.

The $G\alpha_{16}$ subunit couples any activated GPCR to phospholipase C, which results in accumulation of inositol phosphates (IP1), and thus provides a general readout of G-protein activation[22]. Stretch did not
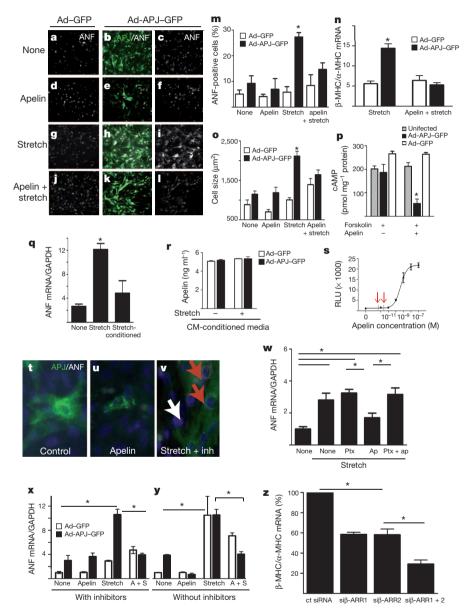

Figure 4 | APJ activation through mechanical stretch elicits cardiac hypertrophy. **a–l**, ANF immunostaining (white) and nuclear DAPI staining (blue) of rat neonatal ventricular cardiomyocytes transduced with rat APJ (Ad-APJ–GFP) or control GFP (Ad-GFP, green). **m**, Quantification of **a–l** ($n = 250$–350 cells). **n**, Quantitative polymerase chain reaction analysis of the ratio between β- and α-MHCs, as an independent index of hypertrophy ($n = 3$–5 samples). **o**, Mean cell sizes in **a–l** ($n = 24$ cells). **p**, Responsiveness of transfected cells to apelin treatment ($n = 4$–5). **q**, ANF expression in cardiomyocytes in the presence of conditioned media from stretched cardiomyocytes ($n = 3$). **r**, Apelin enzyme-linked immunosorbent assay of conditioned media (12 h) from cardiomyocytes (CM) non-stretched and stretched. **s**, Apelin standard curve of CHO APJ–β-arrestin interaction assay; red arrows represent the response elicited by two samples from **r**. **t–v**, Higher

magnification image of neonatal rat ventricular cardiomyocytes showing ANF (white) and APJ (green) expression. White arrow indicates APJ$^-$ cells not expressing ANF. **w**, $G\alpha_i$ inhibition with PTX blocked the ability of apelin to prevent ANF expression ($n = 5$). **x, y**, Quantitative polymerase chain reaction with reverse transcription for ANF in cells treated with (**x**) or without (**y**) inhibitors ($n = 3$). **z**, Diminished expression of hypertrophy markers in Ad-APJ–GFP cardiomyocytes transfected upon knockdown of β-arrestin1 (siβ-ARR1), β-arrestin2 (siβ-ARR2) or both (siβ-ARR1+2) ($n = 4$). Except for **p** and **y**, all experiments were performed in the presence of inhibitors of AT-1 (candesartan, 100 nM) and ET-A (BQ123, 300 nM) added 1 h before stretch and/or apelin treatment until fixation. All are representative experiments performed at least three independent times. Error bars, s.e.m. *$P < 0.05$ between indicated groups, analysis of variance.

increase IP1 production whereas apelin did so in a dose-responsive manner (Fig. 3e). Notably, for any given concentration of apelin, stretch consistently reduced G-protein activation (Fig. 3e). There was a decrease in maximal levels and a shift in half-maximum effective concentration (EC_{50}) from 5.1×10^{-9} to 5.5×10^{-8} when both stimuli were applied simultaneously (Fig. 3e). These experiments demonstrate that APJ activation by stretch is largely G-protein independent and that stretch interferes with apelin-mediated G-protein activation by APJ.

Using a β-arrestin/APJ complementation assay, apelin was found to induce a dose-dependent increase in β-arrestin signalling, as expected (Fig. 3f). Stretch in the absence of apelin boosted β-arrestin signalling twofold more than baseline (Fig. 3g). Notably, stretch increased β-arrestin recruitment in response to apelin at all doses tested (Fig. 3g). Taken together, stretch causes APJ to signal less effectively/potently through G proteins, but to recruit β-arrestin more effectively.

To investigate whether these distinct mechanisms of APJ action differentially affect cardiac hypertrophy, we examined cardiomyocyte cell growth and the expression of molecular markers of pathological hypertrophy in neonatal rat ventricular cardiomyocytes[23,24]. These cardiomyocytes in culture respond to mechanical stretch[25], mainly through the angiotensin II receptor AT-1[26] and the endothelin1 receptor ET-A[27], which are also GPCRs. To test the specific effect of APJ in cardiac stretch/hypertrophy, we used pharmacological inhibitors of AT-1 (100 nM candesartan) and ET-A (300 nM BQ123), hereafter labelled as 'inhibitors'. To overcome APJ down-regulation during cardiomyocyte culture, we re-established APJ expression by adenoviral transduction (about 90% efficiency) with either control GFP (Ad–GFP) or an APJ–GFP fusion protein (Ad-APJ–GFP). Hypertrophy was assessed by the characteristic increase in perinuclear immunolocalization of atrial natriuretic factor (ANF) (Fig. 4a–m). Remarkably, apelin treatment did not increase ANF immunostaining, neither in Ad-APJ–GFP- nor in Ad–GFP-infected cells (Fig. 4d–f, m). In contrast, stretch applied in the presence of inhibitors significantly increased the number of perinuclear ANF+ cells in APJ-restored cardiomyocytes ($27 \pm 1\%$), but not in control cardiomyocytes infected with Ad–GFP ($5.8 \pm 2.1\%$) (Fig. 4g–i, m). Co-stimulation with apelin and stretch reduced the number of ANF+ cells (Fig. 4j–l, m). Stretch also induced specific changes in the expression of other molecular markers of hypertrophy, including an increased ratio of β- to α-myosin heavy chain (β-MHC/α-MHC) (Fig. 4n) and increased cell size (Fig. 4o). Similar to its effects on ANF, the addition of apelin also attenuated effects of stretch on other parameters of hypertrophy (Fig. 4m–o). The induction of ANF by stretch was PTX insensitive, but the ability of apelin to antagonize stretch-induced hypertrophy was prevented by treatment with PTX (Fig. 4w).

Several control experiments confirmed that APJ is directly involved in the response to stretch: neonatal rat ventricular cardiomyocytes treated with forskolin responded to apelin only upon reconstitution of APJ expression, as monitored by their ability to decrease intracellular cAMP levels (Fig. 4p). The possibility that stretch induces the secretion of factors that might indirectly activate hypertrophy through APJ was considered but seems unlikely because conditioned medium from APJ-transfected cardiomyocytes did not induce ANF expression (Fig. 4q). Moreover, the concentration of apelin in the media remained unchanged (approximately $5 \, \text{ng ml}^{-1}$) with or without stretch (Fig. 4r), and conditioned media from stretched cells could not activate β-arrestin recruitment (Fig. 4s). Single-cell analysis of low multiplicity APJ–GFP transduced cells showed that cardiomyocytes require APJ to induce ANF upon stretch (Fig. 4t–v, note perinuclear ANF immuno-staining only in GFP+ cells in panel v), further demonstrating that APJ activation through stretch is sufficient to elicit cardiac hypertrophy. Previous studies showed that APJ interacts with AT-1 and apelin antagonizes AT-1 function[11]. As shown in Fig. 4x, in the presence of AT-1 inhibitors, only those cells expressing APJ responded to stretch by significantly increasing perinuclear ANF expression. In the absence of AT-1 inhibitors, APJ-transduced and non-transduced cells reached

the same maximal level of ANF expression (Fig. 4y). Thus, blocking AT-1 does not impair the ability of APJ to respond to stretch, indicating that APJ alone is sufficient to transduce a stretch-induced hypertrophy signal.

Notably, siRNA specific for β-arrestin1 or β-arrestin2 (Fig. 4z) blocked the stretch induction of hypertrophic markers (β-MHC/α-MHC (Fig. 4z) and ANF (not shown)) with an additive effect when both siRNAs were used together. These data substantiate the model that APJ signalling through β-arrestin mediates stretch-induced myocardial hypertrophy.

In summary, the mechano-response of APJ is necessary (blunted hypertrophic response to TAC of APJ-KO mice, Fig. 1) and sufficient (stretch induction of ANF expression occurs in cells expressing APJ, Fig. 4) to trigger myocardial hypertrophy in a β-arrestin-dependent manner (Fig. 4z). Apelin does not induce hypertrophy, but instead blunts stretch-mediated hypertrophy (Fig. 4j–o), suggesting the ability of apelin to override pathological signalling from stretch. At a mechanistic level, APJ transduces apelin and stretch signalling differently. The response to apelin seems to be G-protein-PTX sensitive, whereas that induced by stretch is PTX insensitive and G-protein independent in the absence of exogenous apelin (Fig. 3). Stretch profoundly affects apelin signalling, diminishing G-protein activation while augmenting β-arrestin recruitment (Fig. 3e–g). These data indicate that APJ integrates apelin and stretch stimuli, biasing the levels of G-protein signalling versus β-arrestin recruitment accordingly.

These results have implications for the consideration of APJ as a drug target, because APJ/stretch can be pathological. Therefore, a beneficial effect will be obtained not by general apelin receptor agonism, but rather by selectively inhibiting the ability of APJ to respond to mechanical stretch or by blocking its interaction with molecules that initiate pathological signalling cascades.

METHODS SUMMARY

All experiments were performed in accordance with relevant guidelines and regulations. The Sanford-Burnham Medical Research Institute's Animal Care & Use Program is accredited by Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC International), and a Multiple Project Assurance A3053-1 is on file in the Office of Laboratory Animal Welfare, Department of Health and Human Services.

APJ and apelin KO mice. APJ-KO mice were obtained from Deltagen. Apelin-KO mice are described elsewhere[13]. Both APJ and apelin lines (male and female) were in C57Bl/6 genetic background in a 99–100% purity, as demonstrated by microsatellite analysis (Radil). For detailed methods, see Supplementary Information.

1. Levy, D., Garrison, R. J., Savage, D. D., Kannel, W. B. & Castelli, W. P. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. *N. Engl. J. Med.* **322,** 1561–1566 (1990).
2. Knowlton, K. U. *et al.* The alpha 1A-adrenergic receptor subtype mediates biochemical, molecular, and morphologic features of cultured myocardial cell hypertrophy. *J. Biol. Chem.* **268,** 15374–15380 (1993).
3. Rapacciuolo, A., Esposito, G., Prasad, S. V. & Rockman, H. A. G protein-coupled receptor signalling in in vivo cardiac overload. *Acta Physiol. Scand.* **173,** 51–57 (2001).
4. Rockman, H. A., Koch, W. J. & Lefkowitz, R. J. Seven-transmembrane-spanning receptors and heart function. *Nature* **415,** 206–212 (2002).
5. O'Dowd, B. F. *et al.* A human gene that shows identity with the gene encoding the angiotensin receptor is located on chromosome 11. *Gene* **136,** 355–360 (1993).
6. Tatemoto, K. *et al.* Isolation and characterization of a novel endogenous peptide ligand for the human APJ receptor. *Biochem. Biophys. Res. Commun.* **251,** 471–476 (1998).
7. Szokodi, I. *et al.* Apelin, the novel endogenous ligand of the orphan receptor APJ, regulates cardiac contractility. *Circ. Res.* **91,** 434–440 (2002).
8. Ashley, E. A. *et al.* The endogenous peptide apelin potently improves cardiac contractility and reduces cardiac loading *in vivo. Cardiovasc. Res.* **65,** 73–82 (2005).
9. Jia, Y. X. *et al.* Apelin protects myocardial injury induced by isoproterenol in rats. *Regul. Pept.* **133,** 147–154 (2006).
10. Siddiquee, K. *et al.* Apelin protects against angiotensin II-induced cardiovascular fibrosis and decreases plasminogen activator inhibitor type-1 production. *J. Hypertens.* **29,** 724–731 (2011).

11. Chun, H. J. *et al.* Apelin signaling antagonizes Ang II effects in mouse models of atherosclerosis. *J. Clin. Invest.* **118,** 3343–3354 (2008).
12. Kuba, K. *et al.* Impaired heart contractility in apelin gene-deficient mice associated with aging and pressure overload. *Circ. Res.* **101,** e32–e42 (2007).
13. Charo, D. N. *et al.* Endogenous regulation of cardiovascular function by apelin-APJ. *Am. J. Physiol. Heart Circ. Physiol.* **297,** H1904–H1913 (2009).
14. Scott, I. C. *et al.* The G protein-coupled receptor Agtrl1b regulates early development of myocardial progenitors. *Dev. Cell* **12,** 403–413 (2007).
15. Frank, D. *et al.* Gene expression pattern in biomechanically stretched cardiomyocytes: evidence for a stretch-specific gene program. *Hypertension* **51,** 309–318 (2008).
16. Zou, Y. *et al.* Mechanical stress activates angiotensin II type 1 receptor without the involvement of angiotensin II. *Nature Cell Biol.* **6,** 499–506 (2004).
17. Iribe, G. *et al.* Axial stretch of rat single ventricular cardiomyocytes causes an acute and transient increase in Ca^{2+} spark rate. *Circ. Res.* **104,** 787–795 (2009).
18. Bollensdorff, C., Lookin, O. & Kohl, P. Assessment of contractility in intact ventricular cardiomyocytes using the dimensionless 'Frank-Starling Gain' index. *Pflügers Arch.* **462,** 39–48 (2011).
19. Civelli, O. GPCR deorphanizations: the novel, the known and the unexpected transmitters. *Trends Pharmacol. Sci.* **26,** 15–19 (2005).
20. Means, C. K., Miyamoto, S., Chun, J. & Brown, J. H. $S1P_1$ receptor localization confers selectivity for G_i-mediated cAMP and contractile responses. *J. Biol. Chem.* **283,** 11954–11963 (2008).
21. Masri, B., Morin, N., Pedebernade, L., Knibiehler, B. & Audigier, Y. The apelin receptor is coupled to G_{i1} or G_{i2} protein and is differentially desensitized by apelin fragments. *J Biol Chem* **281,** 18317–18326 (2006).
22. Kostenis, E. Is $G\alpha_{16}$ the optimal tool for fishing ligands of orphan G-protein-coupled receptors? *Trends Pharmacol. Sci.* **22,** 560–564 (2001).
23. Knowlton, K. U. *et al.* Divergent pathways mediate the induction of ANF transgenes in neonatal and hypertrophic ventricular myocardium. *J. Clin. Invest.* **96,** 1311–1318 (1995).
24. Rockman, H. A. *et al.* Molecular and physiological alterations in murine ventricular dysfunction. *Proc. Natl Acad. Sci. USA* **91,** 2694–2698 (1994).
25. Sadoshima, J. & Izumo, S. Mechanical stretch rapidly activates multiple signal transduction pathways in cardiac myocytes: potential involvement of an autocrine/paracrine mechanism. *EMBO J.* **12,** 1681–1692 (1993).
26. Sadoshima, J., Xu, Y., Slayter, H. S. & Izumo, S. Autocrine release of angiotensin II mediates stretch-induced hypertrophy of cardiac myocytes in vitro. *Cell* **75,** 977–984 (1993).
27. Magga, J., Vuolteenaho, O., Marttila, M. & Ruskoaho, H. Endothelin-1 is involved in stretch-induced early activation of B-type natriuretic peptide gene expression in atrial but not in ventricular myocytes: acute effects of mixed ET(A)/ET(B) and AT1 receptor antagonists *in vivo* and *in vitro*. *Circulation* **96,** 3053–3062 (1997).

'Slings' enable neutrophil rolling at high shear

Prithu Sundd[1], Edgar Gutierrez[2], Ekaterina K. Koltsova[1], Yoshihiro Kuwano[1,3], Satoru Fukuda[4], Maria K. Pospieszalska[1], Alex Groisman[2] & Klaus Ley[1]

Most leukocytes can roll along the walls of venules at low shear stress (1 dyn cm^{-2}), but neutrophils have the ability to roll at tenfold higher shear stress in microvessels *in vivo*[1,2]. The mechanisms involved in this shear-resistant rolling are known to involve cell flattening[3] and pulling of long membrane tethers at the rear[4–6]. Here we show that these long tethers do not retract as postulated[6,7], but instead persist and appear as 'slings' at the front of rolling cells. We demonstrate slings in a model of acute inflammation *in vivo* and on P-selectin *in vitro*, where P-selectin-glycoprotein-ligand-1 (PSGL-1) is found in discrete sticky patches whereas LFA-1 is expressed over the entire length on slings. As neutrophils roll forward, slings wrap around the rolling cells and undergo a step-wise peeling from the P-selectin substrate enabled by the failure of PSGL-1 patches under hydrodynamic forces. The 'step-wise peeling of slings' is distinct from the 'pulling of tethers' reported previously[4–6,8]. Each sling effectively lays out a cell-autonomous adhesive substrate in front of neutrophils rolling at high shear stress during inflammation.

Neutrophil rolling at high shear stress is poorly understood[2]. Some neutrophil adhesion molecules including PSGL-1 are concentrated around the tips of microvilli, where they initiate contact with the vessel wall[9,10]. When the force on a microvillus exceeds a threshold force of ~31 pN, the plasma membrane eventually separates from the cytoskeleton, forming a tether[8,11]. To study the footprints of cells rolling under high shear, we recently developed quantitative dynamic footprinting (qDF)[4] and refined it to allow simultaneous acquisition of two fluorochromes (DqDF)[12]. qDF relies on total internal reflection fluorescence (TIRF) microscopy of fluorescent neutrophils rolling in microfluidic channels on glass substrates coated with P-selectin. We observed that DiI-stained neutrophils isolated from the bone marrow of wild-type mice (Fig. 1a and Supplementary Movie 1) or neutrophils in whole blood of *Lyz2-EGFP* mice[13] (Fig. 1b and Supplementary Movie 2) rolling at 6–10 dyn cm^{-2} frequently formed long processes in the downstream direction. These structures have not been observed previously, and we propose to call them 'slings'. Three-dimensional-reconstructions of the qDF footprint of a rolling neutrophil with a sling in front are shown in Fig. 1c–d (Supplementary Note 1). Scanning electron micrographs (SEMs; Fig. 1e) confirm that neutrophils rolling at 10 dyn cm^{-2} have slings in front and tethers in the rear. As the cells roll forward, the slings are wrapped around the rolling neutrophils, which was revealed both by qDF (Fig. 1f and Supplementary Movie 3) and SEMs (Fig. 1g).

Slings were observed *in vivo* by epifluorescence intravital microscopy of mouse cremaster venules. Slings in front of six different leukocytes (labelled with a monoclonal antibody (mAb) against LFA-1) rolling in cremaster venules of wild-type mice *in vivo* are shown in Fig. 1h–m, Supplementary Fig. 1 and Supplementary Movies 4 and 5.

When a neutrophil is first captured from the flow, an initial sling may take several seconds to form. Slings become more frequent as the rolling progresses at high shear stress (Fig. 2a, b) and their average length increases with shear stress over the range of 6 to 10 dyn cm^{-2} (Fig. 2c; Supplementary Note 2). Because neutrophil rolling at high

shear stress produces three to four tethers per cell[4], we propose that slings may form when detached tethers swing around to the front of the rolling cell. This was confirmed by both qDF (Fig. 2d and Supplementary Movie 6) and differential interference contrast (DIC) microscopy (Fig. 2e and Supplementary Movie 7). Because each tether grows over time, and the aspect ratio of the tether anchorage points suggested increasing force (Supplementary Fig. 2), we calculated the tether force based on the nonlinearly decaying spring viscoelastic (NLDs-viscoelastic) model[8]. Although the anchorage points along newly formed slings initially did not show evidence of deformation, this changed as the cell rolled over the sling, which resulted in the rearmost patch to become load-bearing (white arrowhead at $t = 13$ s in Fig. 1a and $t = 5$ s in Fig. 1b). The force acting on a tether or a load-bearing sling reached a maximum of 80–90 pN just before the tether or sling detached (Fig. 2f). If a rolling cell has two loaded slings and two tethers with the inflection points (white arrowhead at $t = 13$ s in Fig. 1a) and anchorage points (white arrowhead in Supplementary Fig. 2a), respectively, at a distance of 10 μm from the cell centre, then the two slings and two tethers can together balance the entire bond torque and more than three quarters of the forward force acting on the rolling cell (Fig. 2g and Supplementary Fig. 3). Thus, tethers and slings are the major structures which slow down neutrophils rolling at high wall shear stress.

qDF experiments with both naive and Th1 CD4 T cells (Supplementary Note 3; Supplementary Figs 4 and 5) revealed that the formation of slings is indeed associated with stabilized rolling at high shear stress. Next, we addressed the nature of PSGL-1 expression on slings. Raw DqDF images of wild-type neutrophils rolling on P-selectin (Supplementary Fig. 6a) show expression of PSGL-1 (4RB12 mAb[14]; green) on the cell surface, sling and tether anchorage points. Consistent with previous results in suspended cells[9,10], PSGL-1 is localized around the tips of microvilli of rolling neutrophils (Supplementary Fig. 6b), but excluded from the cell surface. As anticipated, PSGL-1 was present in the tether anchorage points, but surprisingly, PSGL-1 was also expressed in patches along the sling. Further experiments were conducted using a red fluorochrome (Alexa-Fluor-568) coupled to 4RB12 mAb (Fig. 3a–c and Supplementary Figs 7 and 8). Specificity was established by investigating PSGL-1-deficient neutrophils rolling on E-selectin which showed no staining for PSGL-1 (Supplementary Fig. 9). We observed a high concentration of PSGL-1 not only at the tip (the former tether anchorage point), but also in discrete patches along the length of slings (Fig. 3a and Supplementary Figs 7 and 8). Analysis of SEM images of slings revealed tips of microvilli preserved as lobes on the slings (Supplementary Fig. 10 and 11). These findings indicate that membrane from several microvilli is pulled from the cell surface to form the tethers. Tethers following detachment become slings and the tips of microvilli, which are rich in PSGL-1, appear as patches on the slings. Indeed, the spacing of PSGL-1 patches on slings (1.6 ± 0.2 (s.e.m.) μm; $n = 34$) is comparable to the estimated distance between the tips of adjacent microvilli (1.32 μm) preserved as lobes on a 10 μm long sling (Supplementary Fig. 12).

[1]Division of Inflammation Biology, La Jolla Institute for Allergy and Immunology, La Jolla, California 92037, USA. [2]Department of Physics, University of California San Diego, La Jolla, California 92093, USA. [3]Department of Dermatology, University of Tokyo, Tokyo 113-8655, Japan. [4]Laboratory of Electron Microscopy, University of Tokyo, Tokyo 113-8655, Japan.
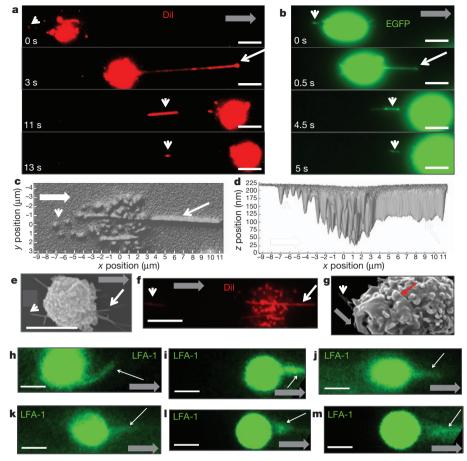
Figure 1 | Rolling neutrophils form slings. a, b, Processed qDF images of a DiI-stained isolated neutrophil (**a**) and an *EGFP* neutrophil in whole blood (**b**) rolling on P-selectin in a microfluidic device. A tether (anchor point marked with white arrowhead at $t = 0$ s) swings around to the front of the cell as a 'sling' (white arrow) at $t = 3$ s (**a**) and 0.5 s (**b**). **c, d,** Three-dimensional reconstruction of the footprint of a rolling neutrophil. **c,** x–y plane and **d,** x–z plane. **e,** SEM of a rolling neutrophil showing tethers and slings. **f,** Unprocessed qDF image of a DiI-stained neutrophil rolling on P-selectin reveals wrapping of slings. **g,** SEM showing sling (red arrow) wrapped around a rolling neutrophil. **h–m,** Epifluorescence intravital microscopic images demonstrating sling formation by six different leukocytes. P-selectin (**a–g**), 20 molecules per μm^2. Shear stress, 10 dyn cm^{-2} (**a, c–g**), 8 dyn cm^{-2} (**b**). Arrowheads, tether anchor points; thin arrow, sling; thick arrow, rolling direction. Scale bars, 5 μm.

As each sling wraps around the rolling cell, each of the PSGL-1 patches functions as a discrete point of attachment with the P-selectin substrate (Fig. 3c, Supplementary Fig. 13 and Supplementary Movie 8). When a PSGL-1 patch on the sling becomes load-bearing, the bond force pulls the patch away from the substrate in the z-direction (Fig. 3d and Supplementary Fig. 14). Concomitantly, the apparent brightness of a PSGL-1 patch decreases by 50–75% (Fig. 3e and Supplementary Fig. 15) as the bonds get loaded because the intensity of the TIRF evanescent wave exponentially decreases with z-distance from the cover slip[4]. On tethers, PSGL-1 is concentrated in the tether anchorage point (Fig. 3b). When the PSGL-1 patch at the tether anchor point breaks, this results in catastrophic failure (step 1 in Fig. 3f), because there is no chance that the bonds can re-form. However, on slings, PSGL-1 is organized in patches and, unlike a tether, a sling peels in a step-wise manner (Fig. 3c, f and Supplementary Fig. 13). Figure 3g shows the impact of the failure of sling patches and tethers on the forward displacement of the cell shown in Fig. 3c. When a patch of PSGL-1 bonds fails, the sling is pulled by the rolling neutrophil until the next PSGL-1 patch becomes loaded. As a consequence, the forward displacement of the rolling cell following a patch failure is smaller (0.2 ± 0.1 (s.d.) μm; $n = 2$; Fig. 3g) than that of a tether failure (1 ± 0.3 (s.d) μm; $n = 4$; Fig. 3g), and the increase in displacement is short lived because a new bond patch is available to be loaded within less than 2 μm. The complete loss of a sling (between 2.2 and 2.4 s in Fig. 3g) is equivalent to the breaking of a tether, which results in a large forward displacement, because no new PSGL-1 patch

is available to bear the load. A similar analysis of two more cells (Supplementary Fig. 16 and 17) show that the slings indeed act as breaks on rolling neutrophils. This was further corroborated by the *in vivo* observation (Supplementary Fig. 18) that rolling leukocytes slow down following sling formation in cremaster venules of mice. 'Step-wise peeling of slings' (white thin arrow in Supplementary Fig. 19) is distinct from 'pulling of tethers'[4,5] (white arrowhead in Supplementary Fig. 19). The two mechanisms probably work synergistically to facilitate shear-resistant rolling.

Under inflammatory conditions *in vivo*, mouse $\alpha_L\beta_2$-integrin LFA-1 binds ICAM-1 (ref. 15) and ICAM-2 (ref. 16) and is known to be responsible for neutrophil slow rolling[17] and arrest[18]. Unlike PSGL-1, LFA-1 is uniformly expressed on the surface of wild-type (Fig. 4a, b, Supplementary Figs 20 and 21, and Supplementary Movie 9) but not LFA-1-deficient neutrophils (Supplementary Fig. 22) and colocalizes with the footprint, tether anchorage points and slings (Fig. 4c). To show directly the differential distribution of PSGL-1 and LFA-1, we labelled the same cell suspension with both non-blocking-anti-PSGL-1 (red or green) and anti-LFA-1 (green or red) mAbs. Again, LFA-1 was seen in the footprint, tether anchorage points and slings, whereas PSGL-1 was restricted to microvilli tips, tether anchorage points and spots on slings (Fig. 4b and Supplementary Figs 23 and 24).

We have previously shown that neutrophils rolling on P-selectin acquire some LFA-1 in the extended conformation with intermediate affinity that can bind to ICAM-1 on the substrate to support a rolling interaction[17,19]. As slings are rich in LFA-1 and wrap around rolling
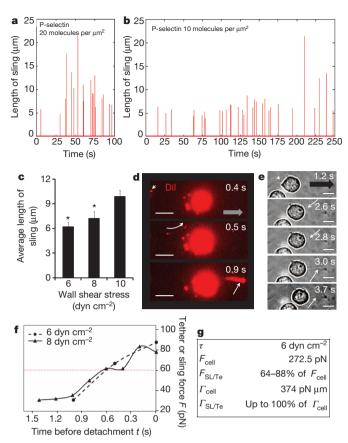
Figure 2 | Sling formation, length and force. a, b, Length of slings formed by a single neutrophil rolling on 20 (**a**) and 10 (**b**) P-selectin molecules per μm^2 in a microfluidic device. **c**, Average length of slings as a function of shear stress. Error bars, s.e.m.; *$P < 0.05$ relative to 10 dyn cm^{-2}; $n = 34$, 47, and 64 for 6, 8, and 10 dyn cm^{-2}, respectively. **d**, A tether (arrowhead at 0.4 s) swings to the front as a sling (white arrow at 0.9 s). qDF images processed. **e**, A tether (arrowhead at 1.2 s) swings to the front as a sling (white arrow at 3 s). DIC images. Shear stress 10 dyn cm^{-2}. P-selectin 20 molecules per μm^2. Scale bars, 5 μm. Rolling direction, thick arrow. **f**, Sling or tether force F before detachment from the substrate. Solid curve, 8 dyn cm^{-2}; dashed curve, 6 dyn cm^{-2}; solid symbols, qDF data. **g**, Fraction of the forward cell force and torque that can be balanced by 2 slings and 2 tethers. τ, shear stress; F_{cell}, total forward force; $F_{SL/Te}$, forward force balanced by 2 slings and 2 tethers; Γ_{cell}, total torque; $\Gamma_{SL/Te}$, torque balanced by 2 slings and 2 tethers. Data in **a–f** representative of five independent experiments.

neutrophils, we investigated whether slings can mediate LFA-1–ICAM-1 or LFA-1–ICAM-2 interactions in *trans* in rolling neutrophils. A recent study has shown that ICAM-1 is upregulated on a subset of mouse whole-blood neutrophils during ischaemia-reperfusion injury[20]. Flow cytometry (Supplementary Fig. 25a–c) showed that ICAM-1 is absent on blood and minimally expressed on bone marrow mouse neutrophils under the conditions used in this study. However, ICAM-2 was abundantly expressed on bone marrow neutrophils (Supplementary Fig. 25d–f). Bone marrow neutrophils from wild-type or LFA-1-deficient (*Itgal*$^{-/-}$) mice either untreated or blocked with function-blocking antibodies against LFA-1 (ref. 17) or ICAM-2 (ref. 21) were allowed to roll on P-selectin and the rolling velocities were compared (Fig. 4d). To our initial surprise, absence of LFA-1 or blocking LFA-1 or ICAM-2 resulted in a small but reproducible decrease in rolling velocity (Fig. 4d). As a positive control, bone marrow neutrophils from wild-type mice were also allowed to roll on glass substrate coated with both P-selectin and ICAM-1. This resulted in much larger decrease in rolling velocity (Supplementary Fig. 26), as shown previously[19]. As expected, treatment with function-blocking anti-LFA-1 mAb returned the rolling velocity to the level seen with LFA-1 blockade on a P-selectin substrate, but not to the higher level
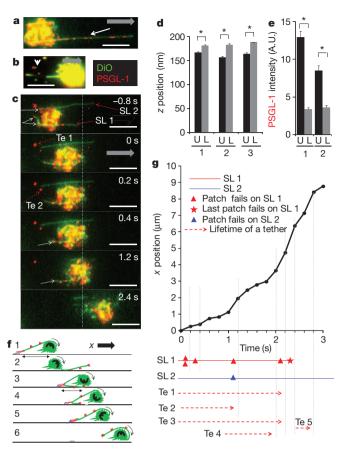

Figure 3 | Neutrophil rolling is stabilized by step-wise peeling of slings. a, PSGL-1 expressed in patches (red spots) on the sling (green; white arrow). **b**, PSGL-1 (red spots) concentrated in the tether anchorage points (white arrowhead). **c**, Two slings and anchorage points of two tethers denoted as SL 1 and SL 2 (red solid arrow) and Te 1 and Te 2 (red dashed arrow), respectively. PSGL-1 patches at the inflection point of SL 1 and SL 2 visible as red spots (solid white arrow) at −0.8 s. As the cell rolls forward the SL 1 and SL 2 become load-bearing at 0 s, which is evident by the apparent loss of PSGL-1 staining (red) in these patches. The two load-bearing patches on SL 1 fail between 0 and 0.2 s and the next two downstream patches on SL 1 become load-bearing (red fluorescence in these patches reduced at 0.2 s). The last PSGL-1 patch on the SL 1 fails between 2.2 and 2.4 s. See Supplementary Fig. 13 for details. **d**, The z-distance of the sling membrane from the cover slip in three different PSGL-1 patches. **e**, PSGL-1 fluorescence intensity for two PSGL-1 patches. U, unloaded. L, loaded. *$P < 0.05$. Error bars, s.e.m. **f**, Schematic of step-wise peeling mechanism. Cell (green), PSGL-1 (red) and P-selectin (blue). **g**, Position of the cell shown in **c** plotted as a function of time. Closed circles, cell position; red and blue solid lines, SL 1 and SL 2, respectively; red and blue closed triangles, failure of individual patches on SL 1 and SL 2, respectively; red five-point star, failure of last patch on SL 1; red-dashed horizontal arrow, lifetime of a tether. Shear stress, 10 dyn cm^{-2}. P-selectin, 20 molecules per μm^2. Scale bars, 5 μm. Rolling direction, thick arrows. DqDF images (**a–c**) were processed. Data representative of six independent experiments.

seen in neutrophils rolling on P-selectin without blockade of LFA-1 or ICAM-2 (Supplementary Fig. 26). DqDF images (Fig. 4e and Supplementary Fig. 27) show that although LFA-1 (red) is expressed both on cell surface and slings, ICAM-2 (green) is expressed mostly on the cell surface. These data indicate that the LFA-1 on slings interacts with ICAM-2 on the cell surface in *trans*. Green staining was absent in neutrophils stained with Alexa-Fluor-488-conjugated rat IgG$_{2aκ}$ (green) antibody (Supplementary Fig. 28). LFA-1 binding to ICAM-2 is expected to cause the sling to be more tightly wrapped around the rolling neutrophil, thus creating a less favourable lever arm (distance d in Fig. 4f). When the LFA-1–ICAM-2 interaction is absent, the lever arm becomes longer (distance d in Fig. 4g) and the rolling velocity decreases. The distance 'd' was indeed found to be significantly larger
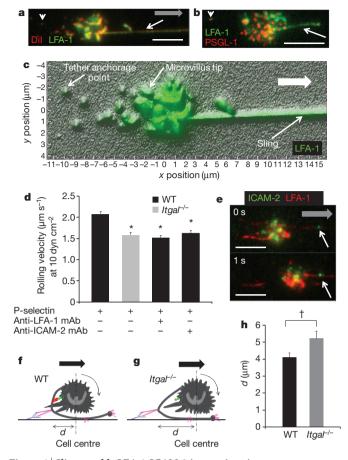
Figure 4 | Slings enable LFA-1-ICAM-2 interactions in *trans*.
a, **b**, Neutrophils stained with DiI (red) and DyLight-488-TIB217 mAb (green) (**a**) or DyLight-488-TIB217 mAb (green) and Alexa-Fluor-568-4RB12 mAb (red) (**b**) were allowed to roll on P-selectin and footprints visualized using DqDF. **c**, The DiI footprint (red; Supplementary Fig. 20) analysed to generate a three-dimensional reconstruction and the LFA-1 image (green; Supplementary Fig. 20) overlaid to reveal distribution of LFA-1 (green) on sling, tether anchorage points, and microvilli (grey hills) versus cell surface (valleys). **d**, Mean rolling velocity of neutrophils on P-selectin. Wild type (WT) (black; $n = 159$), $Itgal^{-/-}$ (grey; $n = 173$), WT blocked with anti-LFA-1 mAb (black; $n = 142$), and WT blocked with anti-ICAM-2 mAb (black; $n = 98$). *$P < 0.01$ relative to WT without blocking. **e**, DqDF images of a rolling neutrophil show distribution of LFA-1 (red) and ICAM-2 (green). **f**, **g**, Schematics showing the mechanism of LFA-1-ICAM-2 interactions in *trans*. *d*, distance of the rearmost load-bearing patch at the inflection point of the sling from the cell centre; LFA-1 (red); ICAM-2 (green); PSGL-1 (violet); P-selectin (blue). **h**, Mean '*d*' plotted for WT ($n = 56$) and $Itgal^{-/-}$ ($n = 53$) neutrophils. †$P < 0.05$. Shear stress, 10 dyn cm^{-2}. P-selectin, 20 molecules per µm^2. Scale bars, 5 µm. White arrowheads, tether anchorage points; white arrows, slings; thick arrow, rolling direction. Error bars, s.e.m. Data representative of five (**a–c**), four (**d**) and three (**e**, **h**) experiments.

in LFA-1-deficient ($Itgal^{-/-}$) neutrophils compared to wild-type neutrophils (Fig. 4h and Supplementary Note 4).

In conclusion, the discovery of slings provides a mechanism by which neutrophils rolling at high shear stress pave their own way for enhanced rolling. The patchy distribution of PSGL-1 along each sling provides a unique adhesive substrate once the cell rolls over the sling. As each PSGL-1 patch fails, a new patch is already lined up that now becomes load-bearing. This step-wise peeling makes slings even more efficient than tethers in slowing down rolling neutrophils. Since slings accumulate over time, eventually, the rolling becomes very stable at high shear stresses[5]. Besides stabilizing rolling, slings are unique structures that also enable rolling neutrophils to present LFA-1 to its ligand ICAM-2 in *trans*. Taken together, catch bonds[22], long tethers[4], cell flattening[1,2,4] and now slings may explain why neutrophils can roll

even at very high shear stress as observed in acute inflammation *in vivo*[1,3].

METHODS SUMMARY

The assembly of microfluidic devices and the coating of cover slips with recombinant murine P-selectin-Fc (20 molecules per µm^2) have been described previously[4,12]. In few experiments, cover slips coated with both recombinant murine P-selectin-Fc and ICAM-1-Fc were used in the microfluidic devices. Neutrophils were isolated from the bone marrow of mice using immune-magnetic negative selection and stained with Vybrant DiO or DiI (Molecular Probes) as described previously[12]. In some experiments, DiO- or DiI-stained neutrophils were additionally stained with Alexa-Fluor-568- or DyLight-488- or Alexa-Fluor-488-conjugated nonblocking-anti-mouse-PSGL-1 (4RB12) or function-blocking anti-mouse-LFA-1 (TIB217) or function-blocking anti-mouse ICAM-2 (3C4) mAb. Cells were suspended in RPMI-1640 (GIBCO-Invitrogen, Carlsbad, CA) + 10% mouse plasma at a concentration of 2.5×10^6 cells per ml and perfused through the microfluidic device at a wall shear stress of 6–10 dyn cm^{-2}. In few experiments, microfluidic devices were perfused with whole blood from anaesthetized *Lyz2-EGFP* mice[13] at a shear stress of 6–10 dyn cm^{-2}. The microfluidic device was placed on the stage of TIRF microscope (IX71; ×100 NA 1.45 Plan-Apochromatic oil immersion TIRFM objective; Olympus America) and images were recorded using qDF[4] or DqDF[12]. qDF images were processed and used to create three-dimensional reconstructions of the footprints as described previously[4,12]. In few experiments, neutrophils rolling in microfluidic devices were fixed with 2% paraformaldehyde and subjected to scanning electron microscopy. Mouse cremaster muscles were prepared for intravital microscopy and 25 µg of DyLight-488-conjugated TIB217 mAb was injected via the catheter connected to the right carotid artery to stain all leukocytes green. A glass cover slip (no. 1; 1 cm × 1 cm) was placed over the moistened cremaster muscle and epifluorescence observations were recorded using a ×100 NA 1.45 oil immersion objective on an Olympus BX61 upright microscope.

Full Methods and any associated references are available in the online version of the paper.

1. Firrell, J. C. & Lipowsky, H. H. Leukocyte margination and deformation in mesenteric venules of rat. *Am. J. Physiol. Heart Circ. Physiol.* **256**, H1667–H1674 (1989).
2. Sundd, P., Pospieszalska, M. K., Cheung, L. S., Konstantopoulos, K. & Ley, K. Biomechanics of leukocyte rolling. *Biorheology* **48**, 1–35 (2011).
3. Damiano, E. R., Westheider, J., Tozeren, A. & Ley, K. Variation in the velocity, deformation, and adhesion energy density of leukocytes rolling within venules. *Circ. Res.* **79**, 1122–1130 (1996).
4. Sundd, P. *et al.* Quantitative dynamic footprinting microscopy reveals mechanisms of neutrophil rolling. *Nature Methods* **7**, 821–824 (2010).
5. Ramachandran, V., Williams, M., Yago, T., Schmidtke, D. W. & McEver, R. P. Dynamic alterations of membrane tethers stabilize leukocyte rolling on P-selectin. *Proc. Natl Acad. Sci. USA* **101**, 13519–13524 (2004).
6. Schmidtke, D. W. & Diamond, S. L. Direct observation of membrane tethers formed during neutrophil attachment to platelets or P-selectin under physiological flow. *J. Cell Biol.* **149**, 719–730 (2000).
7. Shao, J.-Y. in *Leukocyte Adhesion* (ed. Ley, K.) Ch. 2, 25–45 (Academic, 2009).
8. Pospieszalska, M. K., Lasiecka, I. & Ley, K. Cell protrusions and tethers: a unified approach. *Biophys. J.* **100**, 1697–1707 (2011).
9. Hocdé, S. A., Hyrien, O. & Waugh, R. E. Cell adhesion molecule distribution relative to neutrophil surface topography assessed by TIRFM. *Biophys. J.* **97**, 379–387 (2009).
10. Moore, K. L. *et al.* P-selectin glycoprotein ligand-1 mediates rolling of human neutrophils on P-selectin. *J. Cell Biol.* **128**, 661–671 (1995).
11. Waugh, R. E. in *Leukocyte Adhesion* (ed. Ley, K.) Ch. 1, 3–24 (Academic, 2009).
12. Sundd, P. *et al.* Live cell imaging of paxillin in rolling neutrophils by dual-color quantitative dynamic footprinting. *Microcirculation* **18**, 361–372 (2011).
13. Faust, N., Varas, F., Kelly, L. M., Heck, S. & Graf, T. Insertion of enhanced green fluorescent protein into the lysozyme gene creates mice with green fluorescent granulocytes and macrophages. *Blood* **96**, 719–726 (2000).
14. Pendl, G. G. *et al.* Immature mouse dendritic cells enter inflamed tissue, a process that requires E- and P-selectin, but not P-selectin glycoprotein ligand 1. *Blood* **99**, 946–956 (2002).
15. Springer, T. A. Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm. *Cell* **76**, 301–314 (1994).
16. Xu, H. *et al.* Characterization of murine intercellular adhesion molecule-2. *J. Immunol.* **156**, 4909–4914 (1996).
17. Zarbock, A., Lowell, C. A. & Ley, K. Spleen tyrosine kinase Syk is necessary for E-selectin-induced $\alpha_L\beta_2$ integrin-mediated rolling on intercellular adhesion molecule-1. *Immunity* **26**, 773–783 (2007).

18. Lawrence, M. B. & Springer, T. A. Leukocytes roll on a selectin at physiological flow rates: distinction from and prerequisite for adhesion through integrins. *Cell* **65,** 859–873 (1991).
19. Kuwano, Y., Spelten, O., Zhang, H., Ley, K. & Zarbock, A. Rolling on E- or P-selectin induces the extended but not high-affinity conformation of LFA-1 in neutrophils. *Blood* **116,** 617–624 (2010).
20. Woodfin, A. *et al.* The junctional adhesion molecule JAM-C regulates polarized transendothelial migration of neutrophils *in vivo. Nature Immunol.* **12,** 761–769 (2011).
21. Huang, M. T. *et al.* ICAM-2 mediates neutrophil transmigration *in vivo*: evidence for stimulus specificity and a role in PECAM-1-independent transmigration. *Blood* **107,** 4721–4727 (2006).
22. Marshall, B. T. *et al.* Direct observation of catch bonds involving cell-adhesion molecules. *Nature* **423,** 190–193 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.L. (klaus@liai.org).

METHODS

Reagents. Recombinant murine P-selectin-Fc, E-selectin-Fc and ICAM-1-Fc were purchased from R&D Systems. Function-blocking mAbs RB40.34 (rat IgG$_1$) against mouse P-selectin, TIB217 (rat IgG$_{2a}$) against mouse LFA-1, 4RA10 (rat IgG$_1$) against mouse PSGL-1, and 2.4G2 (rat IgG$_{2b}$) against mouse Fc-receptor were purified from hybridoma supernatant at the biomolecular facility of the University of Virginia, Charlottesville, Virginia, USA. Function-blocking antibody 3C4 (rat IgG$_{2aκ}$) against mouse ICAM-2 was purchased from BD Pharmingen. Non-blocking mAb 4RB12 (rat IgG$_{2a}$) against mouse PSGL-1 was a gift from D. Vestweber. Alexa-Fluor-488-conjugated mAb 3C4 (rat IgG$_{2aκ}$) against mouse ICAM-2 and Alexa-Fluor-488-conjugated rat IgG$_{2aκ}$ isotype control antibody were purchased from BioLegend. Fluorescein-isothiocyanate (FITC) conjugated YN1 (rat IgG$_{2bκ}$) mAb against mouse ICAM-1 and FITC-conjugated rat IgG$_{2bκ}$ isotype control antibody were purchased from eBioscience. DyLight-488 mAb labelling kit (catalogue no. 53025) was purchased from Thermo Fisher Scientific. Alexa-Fluor-568 mAb labelling kit (catalogue no. A-20184), Vybrant DiI (catalogue no. V-22885) and DiO (catalogue no. V-22886) were purchased from Molecular Probes.

Mice. Wild-type (C57BL/6J) mice (6–8 weeks old) were purchased from the Jackson Laboratory. *Lyz2-EGFP* mice with the *EGFP* gene inserted into the lysozyme M (LysM, gene name, *Lyz2*) locus have been shown to express *EGFP* in the cytosol specifically in cells of myelomonocytic lineage (granulocytes and macrophages)[13]. *Lyz2-EGFP*, PSGL-1-deficient (*Selplg$^{−/−}$*)[23], and LFA-1-deficient (*Itgal$^{−/−}$*)[24] mice on C57BL/6J background were bred in-house in a pathogen-free animal facility at the La Jolla Institute for Allergy and Immunology (LIAI). Mice were handled according to the guidelines set by the Department of Laboratory Animal Care (DLAC) at LIAI and all surgical procedures were done as per the guidelines in the protocol approved by the Animal Care Committee of LIAI.

Microfluidic device. The assembly of microfluidic devices used in this study and the coating of cover slips with recombinant murine P-selectin-Fc have been described previously[4,12]. P-selectin coating concentration of 2 and 1 µg ml$^{−1}$ resulted in a P-selectin molecular density of ∼20 and 10 molecules per µm^2, respectively, which is comparable to the P-selectin molecular density observed on endothelial cells[25,26]. In some experiments (Supplementary Fig. 26), coverslips were coated with a cocktail containing 2 µg ml$^{−1}$ of recombinant murine P-selectin-Fc and 10 µg ml$^{−1}$ of recombinant murine ICAM-1-Fc to create a substrate expressing both P-selectin and ICAM-1. Radioimmunoassay confirmed that the P-selectin molecular density was unaltered (20 molecules per µm^2) by the presence of ICAM-1 in the coating cocktail (data not shown).

Microfluidic perfusion assay. Neutrophils were isolated from the bone marrow of mice using immune-magnetic negative selection (EasySep mouse neutrophil enrichment kit; catalogue no. 19709A; STEMCELL Technologies) and stained with an intercalating membrane dye Vybrant DiO or DiI as described previously[12]. In some experiments, DiO- or DiI-stained neutrophils were additionally stained with Alexa-Fluor-568- or DyLight-488-conjugated 4RB12 or TIB217 mAbs. Stained (DiO or DiI) or unstained neutrophils (5–10 × 10^6 cells) were suspended in 200 µl Hanks Balanced Salt Solution without Ca^{2+} and Mg^{2+} (HBSS; GIBCO-Invitrogen) + 1% bovine serum albumin (BSA; catalogue no. A7906, Sigma-Aldrich), pH 7.4 and 2 µl of Fc-receptor blocking mAb 2.4G2 (0.5 mg ml$^{−1}$) was added to the cell suspension and cells were incubated for 10 min at room temperature. After 10 min, 4 µl of DyLight-488- or Alexa-Fluor-568-conjugated TIB217 mAb (0.62 mg ml$^{−1}$) or 2 µl of Alexa-Fluor-568- or DyLight-488-conjugated 4RB12 mAb (0.6 mg ml$^{−1}$) or both were added to the cell suspension and the cells were incubated for 30 min at room temperature. After 30 min, cells were washed twice in HBSS + 1% BSA and then, once in RPMI-1640 (GIBCO-Invitrogen) in a centrifuge (Sorvall RT, Thermo Scientific) at 300*g* (5 min, 22 °C). Cells were finally suspended in RPMI-1640 + 10% mouse plasma at a concentration of 2.5 × 10^6 cells per ml and perfused through the microfluidic device at a wall shear stress of 6–10 dyn cm$^{−2}$. The microfluidic device was placed on the stage of the TIRF microscope and images were recorded using qDF[4] or DqDF[12]. In a different set of experiments (Fig. 4d and Supplementary Fig. 26), bone marrow neutrophils from wild-type or *Itgal$^{−/−}$* mice (2.5 × 10^6 cells per ml in RPMI + 10% mouse plasma) either untreated or blocked with 30 µg ml$^{−1}$ (in cell suspension) of function-blocking TIB217 or 3C4 mAb were perfused through the microfluidic device at a wall shear stress of 10 dyn cm$^{−2}$. The data shown in Fig. 4e and Supplementary Figs 27 and 28 was generated by incubating bone marrow neutrophils from wild-type mice with 2 µl of 2.4G2 mAb (0.5 mg ml$^{−1}$) in 200 µl of HBSS + 1% BSA at room temperature for 10 min. After 10 min, 4 µl of Alexa-Fluor-568-conjugated TIB217 mAb (0.62 mg ml$^{−1}$) together with 4 µl of Alexa-Fluor-488-conjugated 3C4 mAb (0.5 mg ml$^{−1}$) or rat IgG$_{2aκ}$ isotype control Ab (0.5 mg ml$^{−1}$) was added and cells were incubated for 30 min at room temperature. Finally, cells were washed twice with HBSS + 1% BSA and once with RPMI-1640 in a centrifuge at 300*g* (5 min, 22 °C) and suspended in RPMI-1640 + 10% mouse plasma at a concentration of 2.5 × 10^6 cells per ml. Cells were perfused through the microfluidic device at a wall shear stress of 10 dyn cm$^{−2}$ and images were recorded using DqDF. Incubation of neutrophils with function blocking anti-mouse PSGL-1 (4RA10) mAb abolished all rolling on P-selectin.

Auto-perfused microfluidic perfusion assay. The auto-perfused microfluidic perfusion assay has been described previously[4] in detail.

Isolation of naive CD4 T cells and differentiation into Th1 cells. Naive CD4 T cells were isolated from the spleen of wild-type mice using EasySep-mouse-CD4T-cell enrichment kit (STEMCELL Technologies) and *in vitro* differentiated into Th1 cells as described elsewhere[27]. Cells were stained with 1.5 µM DiI and suspended in RPMI-1640 + 10% mouse plasma at a concentration of 2.5 × 10^6 cells per ml. Incubation of Th1 cells with function-blocking anti-mouse PSGL-1 (4RA10) mAb abolished all rolling on P-selectin.

Expression of ICAM-1 and ICAM-2 on mouse neutrophils. Expression of ICAM-1 and ICAM-2 on mouse blood and bone marrow neutrophils were measured using flow cytometry. Blood and bone marrow neutrophils from wild-type mice were stained with FITC-conjugated YN1 mAb against mouse ICAM-1- or Alexa-Fluor-488-conjugated 3C4 mAb against mouse ICAM-2 or respective isotype control Abs and flow cytometric analysis was done using BD FACSCanto II flow cytometer, BD Biosciences.

qDF and DqDF microscopy. The qDF and DqDF setups and the theory of qDF have been described previously in detail[4,12]. The set up consisted of an IX71 inverted TIRF research microscope (Olympus America) with a ×100 NA 1.45 Plan-Apochromatic oil immersion TIRFM objective and 30 mW blue ($λ = 488$ nm) and 20 mW yellow-green ($λ = 561$ nm) diode-pumped-solid-state lasers (CVI Melles Griot) as TIRF excitation light sources. Images were captured at a rate of 2 to 8 frames per s using DV2 (Photometrics) DualView video coupler and a 16-bit digital CCD camera (Hamamatsu C10600-10B ORCA-R^2). The laser shutters and camera were controlled with the SlideBook5.0 software (Intelligent Imaging Innovations), which has built-in features for image analysis. The absorption and emission peaks of the fluorochromes used in this study were, respectively, 488 and 507 nm for *EGFP*, 484 and 501 nm for DiO, 549 and 565 nm for DiI, 493 and 518 nm for DyLight-488, 579 and 603 nm for Alexa-Fluor-568, and 495 and 515 nm for Alex-Fluor-488. A TIRF incidence angle of $θ = 70°$ was used for both lasers in all qDF and DqDF experiments.

Image processing of qDF images. qDF and DqDF images were saturated to reveal slings and tether anchorage points as described previously[4]. Briefly, the scale of the intensity histogram of a qDF image was flattened by fixing the maxima close to the minima to reveal fine structures like slings and tether anchor points (Figs 1a, b, 2d and 3a–c). Saturation of images causes cells to appear bigger than their actual size.

Creation of three-dimensional reconstructions and two-dimensional colour maps. Raw DiO or DiI qDF images were used to create three-dimensional reconstructions (Figs 1c, d and 4c and Supplementary Figs 6b and 7b) or two-dimensional colour maps (Supplementary Fig. 14a) of the footprints as described previously[4,12].

Scanning electron microscopy. Neutrophils from wild-type mice rolling on P-selectin (20 molecules per µm^2) in a microfluidic device at a shear stress of 10 dyn cm$^{−2}$ were fixed by perfusing a cocktail containing 2% paraformaldehyde, 2.5% glutaraldehyde, 1 mM CaCl$_2$ and 0.1 M sodium cacodylate buffer (pH 7.4) in distilled water. Fixed cells were visualized using a scanning electron microscope (S-450; Hitachi) as described elsewhere[28]. The diameter of the sling and neutrophil were estimated from electron micrographs using NIH-Image J (National Institutes of Health, Bethesda, Maryland, USA). The reported values are corrected based on 30% shrinkage as determined by the cell diameter measurements before and after fixation.

Epifluorescence intravital microscopy. Surgical-trauma-induced rolling of neutrophils in the cremaster venules of mice is a model of acute inflammation[29] and the rolling within 60 min following surgery is primarily mediated by P-selectin on the endothelium binding to PSGL-1 on neutrophils[30,31]. Each wild-type mouse was anaesthetized, its trachea was intubated, right carotid artery was cannulated, and the cremaster muscle was prepared for intravital microscopy as described previously[29]. To increase contrast, neutrophil surfaces were labelled *in vivo* with DyLight-488-conjugated TIB217 mAb (green) against mouse LFA-1, which is highly expressed on the cell surface including slings (Fig. 4c). Immediately before microscopic observations, 25 µg of DyLight-488-conjugated TIB217 mAb was injected via the catheter connected to the right carotid artery. A glass cover slip (no. 1; 1 cm × 1 cm) was placed over the moistened cremaster muscle and epifluorescence observations were recorded using a ×100 NA = 1.45 Plan-Apochromatic oil immersion objective on an Olympus BX61 upright microscope. Images were recorded using a 16-bit digital CCD camera (Hamamatsu C10600-10B ORCA-R^2) which was connected to a PC (Dell Precision T3400; Intel Core 2 Duo CPU-3.16 GHz; 3.25 GB RAM) through an IEEE1394b interface. Images were

recorded at the rate of 7–16 frames s^{-1} and 2×2 binning as well as contrast enhancement (Fig. 1h–m, Supplementary Fig. 1 and Supplementary Movies 4 and 5) was done post-acquisition using SlideBook5.0.

Estimation of tether and sling force. At any of the steps 3–5 during the step-wise peeling of a sling (Fig. 3f), the sling acts just like a tether that can balance not only the shear force but also the torque acting on the rolling cell. Thus, a sling was treated as a tether in estimating its contribution in sharing of hydrodynamic load (Fig. 2f). Figure 2f shows the force F acting on a tether or sling starting from the time point when F exceeds 31 pN for the first time and before the tether or the load-bearing patch on the sling detaches from the substrate ($t = 0$ s; step 6 in Fig. 3f). The parameters used for force calculation were the cell radius[32] $r_c = 4$ μm, membrane curvature modulus[11] $k_c = 0.2$ pN μm, membrane interfacial drag coefficient[8] $d_{int} = 600$ pN s μm^{-3}, tether threshold force[8] $F_{th} = 31$ pN, and average tether anchorage point-to-substrate distance of 135 nm (based on qDF measurements). The variables used for the calculation of force, namely cell translational velocity V_{tr}, cell–substrate separation distance, and the horizontal distance between the tether anchorage point and the projection of the cell centre on the substrate (the anchor-cell-centre projection line) were measured experimentally using qDF. However, the exact location of the base of a tethered microvillus cannot be determined precisely by qDF measurements. For simplicity, we assumed that the projection of the tether's microvillus base is on the anchor-cell-centre projection line when the tether anchorage point starts to be seen (changing reasonably the location of the microvillus base yielded similar results). The tether extension rate dL/dt, where L is the tether extension, was calculated geometrically assuming that the cell rotational velocity[33] $V_{rot} = 0.97 V_{tr}$. The tether force F was calculated numerically from the formula $dL/dt = (F^3 - F(F_{th})^2)/(16\pi^3(k_c)^2 d_{int}\ln(r_c F/2\pi k_c))$, assuming the tether to be a nonlinearly decaying spring viscoelastic (NLDs-viscoelastic)[8,34].

Fraction of cell forward force and torque balanced by slings and tethers. On the basis of the findings of the current and previous study[4], a plausible scenario was considered where a rolling neutrophil has two peeling slings, each with the inflection point or the loaded patch at a distance of 10 μm from the cell centre (distance shown by double-headed arrow in step 3 of Fig. 3f) and two tethers with anchorage points at a distance of 10 μm from the cell centre (distance shown by double headed arrow in step 1 of Fig. 3f). The total forward force and torque acting on the rolling cell were estimated for wall shear stresses of 6 and 8 dyn cm^{-2} using Goldman's theory[35–38], assuming that the cell is spherical of radius $r_c = 3.75$ μm, the translational velocity $V_{tr} = 2.4$ μm s^{-1} (based on qDF data), and the rotational velocity[33] is $0.97 V_{tr}$. The fractions of forward force and torque that can be balanced by two slings and two tethers are shown in Fig. 2g and Supplementary Fig. 3. The values were calculated based on geometry of the tethers or slings having projection lengths of 10 μm and force loads of 60 pN (average force shown by red line in Fig. 2f). The fractions were also calculated for $r_c = 4$ μm (data not shown).

Statistical analysis. The pairs of means for the data sets in Figs 3d, e and 4h and Supplementary Figs 14 and 15 were compared, each separately, using student t-tests with equal variances. The pairs of means for the data sets in Figs 2c, 4d and Supplementary Fig. 26 were compared using Bonferroni t-tests[39] where the significance levels for individual pairs were adjusted based on the number of data entries and number of comparison groups to secure the overall significance level of 0.05.

23. Xia, L. *et al.* P-selectin glycoprotein ligand-1-deficient mice have impaired leukocyte tethering to E-selectin under flow. *J. Clin. Invest.* **109**, 939–950 (2002).
24. Ding, Z. M. *et al.* Relative contribution of LFA-1 and Mac-1 to neutrophil adhesion and migration. *J. Immunol.* **163**, 5029–5038 (1999).
25. Hattori, R., Hamilton, K., Fugate, R., McEver, R. & Sims, P. Stimulated secretion of endothelial von Willebrand factor is accompanied by rapid redistribution to the cell surface of the intracellular granule membrane protein GMP-140. *J. Biol. Chem.* **264**, 7768–7771 (1989).
26. Doyle, E. L. *et al.* CD63 is an essential cofactor to leukocyte recruitment by endothelial P-selectin. *Blood* **118**, 4265–4273 (2011).
27. Atarashi, K., Hirata, T., Matsumoto, M., Kanemitsu, N. & Miyasaka, M. Rolling of Th1 cells via P-selectin glycoprotein ligand-1 stimulates LFA-1-mediated cell binding to ICAM-1. *J. Immunol.* **174**, 1424–1432 (2005).
28. Tojo, A. *et al.* Nitric oxide generated by nNOS in the macula densa regulates the afferent arteriolar diameter in rat kidney. *Med. Electron Microsc.* **37**, 236–241 (2004).
29. Ley, K. *et al.* Sequential contribution of L- and P-selectin to leukocyte rolling in vivo. *J. Exp. Med.* **181**, 669–675 (1995).
30. Kunkel, E. J. *et al.* Absence of trauma-induced leukocyte rolling in mice deficient in both P-selectin and intercellular adhesion molecule 1. *J. Exp. Med.* **183**, 57–65 (1996).
31. Sperandio, M. *et al.* Severe impairment of leukocyte rolling in venules of core 2 glucosaminyltransferase-deficient mice. *Blood* **97**, 3812–3819 (2001).
32. Schmid-Schönbein, G. W. Leukocyte biophysics. An invited review. *Cell Biophys.* **17**, 107–135 (1990).
33. Pospieszalska, M. K. & Ley, K. in *Leukocyte Adhesion* (ed. Ley, K.) Ch. 8 221–296 (Academic, 2009).
34. Brochard-Wyart, F., Borghi, N., Cuvelier, D. & Nassoy, P. Hydrodynamic narrowing of tubes extruded from cells. *Proc. Natl Acad. Sci. USA* **103**, 7660–7663 (2006).
35. Goldman, A. J., Cox, R. G. & Brenner, H. Slow viscous motion of a sphere parallel to a plane wall. II. Couette flow. *Chem. Eng. Sci.* **22**, 653–660 (1967).
36. Pospieszalska, M. K., Zarbock, A., Pickard, J. E. & Ley, K. Event-tracking model of adhesion identifies load-bearing bonds in rolling leukocytes. *Microcirculation* **16**, 115–130 (2009).
37. Krasik, E. F. & Hammer, D. A. A semianalytic model of leukocyte rolling. *Biophys. J.* **87**, 2919–2930 (2004).
38. Goldman, A. J., Cox, R. G. & Brenner, H. Slow viscous motion of a sphere parallel to a plane wall. I. Motion through a quiescent fluid. *Chem. Eng. Sci.* **22**, 637–651 (1967).
39. Wilcox, R. R. *New Statistical Procedures For The Social Sciences: Modern Solutions To Basic Problems* (Lawrence Erlbaum Associates, 1987).

LETTER

A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response

Laurens Kruidenier[1]*, Chun-wa Chung[2]*, Zhongjun Cheng[3]*, John Liddle[1]*, KaHing Che[4,5]*, Gerard Joberty[6], Marcus Bantscheff[6], Chas Bountra[4], Angela Bridges[2], Hawa Diallo[1], Dirk Eberhard[6], Sue Hutchinson[2], Emma Jones[2], Roy Katso[2], Melanie Leveridge[2], Palwinder K. Mander[1], Julie Mosley[2], Cesar Ramirez-Molina[1], Paul Rowland[2], Christopher J. Schofield[4], Robert J. Sheppard[1], Julia E. Smith[1], Catherine Swales[5], Robert Tanner[2], Pamela Thomas[2], Anthony Tumber[4], Gerard Drewes[6], Udo Oppermann[4,5], Dinshaw J. Patel[3], Kevin Lee[1]† & David M. Wilson[1]

The jumonji (JMJ) family of histone demethylases are Fe^{2+}- and α-ketoglutarate-dependent oxygenases that are essential components of regulatory transcriptional chromatin complexes[1–4]. These enzymes demethylate lysine residues in histones in a methylation-state and sequence-specific context[5]. Considerable effort has been devoted to gaining a mechanistic understanding of the roles of histone lysine demethylases in eukaryotic transcription, genome integrity and epigenetic inheritance[2,4,6], as well as in development, physiology and disease[3,7]. However, because of the absence of any selective inhibitors, the relevance of the demethylase activity of JMJ enzymes in regulating cellular responses remains poorly understood. Here we present a structure-guided small-molecule and chemoproteomics approach to elucidating the functional role of the H3K27me3-specific demethylase subfamily (KDM6 subfamily members JMJD3 and UTX)[8]. The liganded structures of human and mouse JMJD3 provide novel insight into the specificity determinants for cofactor, substrate and inhibitor recognition by the KDM6 subfamily of demethylases. We exploited these structural features to generate the first small-molecule catalytic site inhibitor that is selective for the H3K27me3-specific JMJ subfamily. We demonstrate that this inhibitor binds in a novel manner and reduces lipopolysaccharide-induced proinflammatory cytokine production by human primary macrophages, a process that depends on both JMJD3 and UTX. Our results resolve the ambiguity associated with the catalytic function of H3K27-specific JMJs in regulating disease-relevant inflammatory responses and provide encouragement for designing small-molecule inhibitors to allow selective pharmacological intervention across the JMJ family.

Initially, we applied a structure–function approach to elucidate the underlying recognition principles that account for the exquisite specificity of JMJD3 for H3K27me3 (trimethylated lysine 27 on histone 3) over other histone peptides containing the common ARKS amino acid motif[9]. We solved high-resolution crystal structures of the approximately 500 carboxy-terminal amino acids of mouse and human JMJD3 in the presence of cofactor and metal (Fig. 1a, Supplementary Fig. 1a and Supplementary Tables 1 and 2). The orthologue structures superimpose well on each other. The crystal structure of the histone H3(20–34)K27me3 peptide bound to an engineered mouse JMJD3 construct in the presence of the cofactor analogue N-oxalylglycine (NOG) and metal (Ni^{2+}) at 2.5 Å resolution is shown in Fig. 1b. This truncated JMJD3 protein encompasses residues 1,157 to 1,641 and includes a cleaved insertion segment that replaces a disordered loop, to facilitate crystal packing and peptide binding (Supplementary Fig. 2a, b).
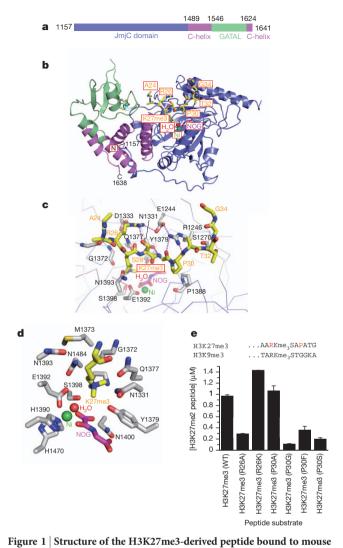
The JMJD3 enzyme is composed of a JmjC catalytic domain and a C-terminal segment within which is embedded a Zn^{2+}-coordinated GATA-like (GATAL) domain of novel topology that is flanked by α-helical segments that form a four-helix bundle (Fig. 1b and Supplementary Fig. 1b, c). The two domains pack against each other with a large buried surface area (4,115 Å^2), explaining the requirement for the C-terminal domain for optimal stability and catalytic competence of the truncated JMJD3 protein (Supplementary Fig. 3).

The H3K27me3 peptide binds in a channel within the catalytic domain and amino acid residues A24 to G34 make several hydrogen-bonding and van der Waals interactions with the enzyme as is evident in Fig. 1c (see also Supplementary Fig. 4). The K27me3 group inserts deep into the catalytic pocket, bringing it close to the NOG and Ni^{2+} that lie at its base (Fig. 1d). The hydrogen-bonding interactions that anchor the H3K27 residue are shown in Supplementary Fig. 5a. The methylene side chain of K27me3 forms hydrophobic contacts with M1373. Elsewhere, the catalytic pocket is lined with polar residues: asparagine, glutamine, serine and tyrosine. The structure shows that the histone peptide residues both amino- and C-terminal to K27, especially P30 and R26, make several interactions (such as with R1246, P1388 and E1244 of JMJD3) that are important for specificity and recognition.

Both H3K9 and H3K27 have a common ARKS sequence context. It has been shown that H3K9me2 and H3K27me3 peptides bind with the same directionality to KDM7A, a dual-specificity H3K9/H3K27 demethylase from *Caenorhabditis elegans*[9] (Supplementary Fig. 6a, b). The present study shows that the H3K27me3 peptide binds to the H3K27-specific demethylase JMJD3 in the opposite orientation to KDM7A. In part, this difference is associated with P30, which points inwards in the complex with JMJD3 but outwards in the complex with KDM7A (Supplementary Fig. 6b, c). In addition, a comparison of intermolecular contacts in our H3K27me3–JMJD3 structure with the corresponding structures of H3K9me2–PHF8 (ref. 10) and H3K9me3–JMJD2A[11,12] establishes that residues outside the ARKS segment also contribute to substrate specificity (Supplementary Fig. 7).

To understand the substrate recognition and to determine the impact on lysine demethylase activity, we mutated key residues of the H3K27me3 peptide–enzyme complex, as determined from the crystal structure (Supplementary Figs 8 and 9). In the bound state of the H3 peptide, R26 forms hydrogen bonds directly with E1244 and D1333 (Fig. 1c). The importance of these hydrogen bonds is shown by the lack of substrate turnover of the R26A mutant peptide, in which the arginine at position 26 has been replaced by alanine (Fig. 1e and Supplementary Fig. 8). Mutation of the P30 residue of H3, which is directed towards P1388 of JMJD3 (Fig. 1c), to glycine, serine or phenylalanine resulted in a loss of activity (Fig. 1e). The P30A mutant

Figure 1 | Structure of the H3K27me3-derived peptide bound to mouse JMJD3. a, Domain topology and numbering system of JMJD3 construct used for crystallization of the enzyme in the free state. **b**, Overview of the structure of H3(20–34)K27me3–mouse JMJD3(1157–1641) complex, with the histone peptide (K27me3, yellow) in a stick representation and the engineered enzyme (Supplementary Fig. 2) colour coded by domain as in **a**: GATAL (green), JmjC (blue) and helical (pink) domains. The heteroatoms in the stick representation are coloured according to atom type: oxygen (red) and nitrogen (blue). The N and C termini of the enzyme are indicated. **c**, View of the intermolecular interactions between the bound histone peptide (yellow) and the enzyme residues (grey), with hydrogen bonds shown as dashed pink lines. **d**, View of the catalytic pocket of the enzyme (grey), the base of which is lined by NOG (pink) and Ni^{2+} (green). The histone peptide is shown in yellow. Details of the hydrogen bonding in the catalytic pocket are outlined in Supplementary Fig. 5a. **e**, Effect of mutations of the native H3K27me3 substrate peptide (highlighted in red) at R26 and P30 on turnover. Mass spectrometric quantification of the H3K27me2 product was performed after a 6-min reaction with concentrations of peptide spanning 0.6 to 250 μM (Supplementary Fig. 8). Data are shown at approximately the K_m (Michaelis constant) of the standard, wild-type (WT), H3K27me3 peptide (22 μM) for clarity. Data are presented as the mean ± s.d.

retained activity, highlighting the requirement for an appropriate hydrophobic interaction with P1388 (Fig. 1e). It has been observed that serine/threonine kinases such as MSK1 can phosphorylate H3S28 (ref. 13). Interestingly, JMJD3 did not demethylate the H3K27me3 peptide when it was phosphorylated at S28 (H3K27me3S28ph) (Supplementary Fig. 9). Our findings also indicate that the JMJD3 residues E1244, R1246, D1333 and P1388 are likely to be important for substrate recognition, as single mutations of these residues to alanine caused a loss of lysine demethylation activity. Structural and sequence comparisons of JMJD3, UTX and UTY suggest that these

recognition determinants are preserved within the KDM6 subfamily of demethylases[14] (Supplementary Figs 10–12).

We exploited these structural insights, in particular the protein interactions made by α-ketoglutarate and the P30 residue of the H3 peptide, to optimize a series of weakly active hits discovered from a screening of the GlaxoSmithKline corporate compound collection (~2 million compounds[15]). The resultant lead, GSK-J1, had a half-maximum inhibitory concentration (IC_{50}) of 60 nM in the JMJD3 AlphaScreen assay (Fig. 2a). A 1.8 Å co-crystal structure of GSK-J1 bound to human JMJD3 revealed the critical interactions within the active site (Fig. 2b, Supplementary Table 2 and Supplementary Fig. 13 (omit electron-density map)). The propanoic acid of GSK-J1 mimics α-ketoglutarate binding by maintaining interactions with K1381, T1387 and N1480 (Fig. 2b and Supplementary Fig. 14a). The aromatic ring of the tetrahydrobenzazepine of GSK-J1 sits in a narrow cleft between R1246 and P1388 (Fig. 2b and Supplementary Fig. 14b), mimicking P30 of the histone peptide. GSK-J1 is competitive with α-ketoglutarate but non-competitive with the peptide substrate. This finding suggests that GSK-J1's mechanism does not follow a simple random model for bisubstrate inhibition[16] (Supplementary Fig. 15).

The pyridyl-pyrimidine biaryl of GSK-J1 makes a bidentate interaction with the catalytic metal (Fig. 2b) and induces a shift in the Co^{2+} ion (used in this crystallization experiment to mimic the Fe^{2+} ion) of 2.34 Å away from the conserved HHE triad (Supplementary Fig. 16). Uniquely, this movement results in the metal cation swapping position with a previously apical water molecule. Subsequently, H1470 makes an indirect water-bridged interaction with the metal, whereas the direct interactions of the metal with H1390 and E1392 are preserved through small changes in their side-chain conformations (Fig. 2c and Supplementary Fig. 14c). The discovery of this subtle but distinctive dynamic metal shift in JMJD3 provides a further opportunity to derive selectivity by designing compounds that take advantage of this elasticity (Supplementary Fig. 17 and Supplementary Information (Materials)).

The overlay of the peptide and inhibitor complexes (Fig. 2c and Supplementary Fig. 18) highlights the shift in the divalent cation, as well as the overlap between bound GSK-J1 and the NOG binding site, thereby accounting for cofactor displacement in the inhibitor complex and the observed mechanism of action.

GSK-J1 is selective for the H3K27 demethylases JMJD3 and UTX and is inactive against a panel of demethylases of the JMJ family, as measured by a combination of thermal shift, mass spectrometry[17] and antibody-based assays[18] (Fig. 3a, b, Supplementary Fig. 19 and Supplementary Table 3). GSK-J1 also did not significantly inhibit 100 protein kinases at a concentration of 30 μM in a competition-binding assay (Supplementary Table 4) and had negligible off-target activity against a panel of 60 unrelated proteins, including other chromatin-modifying enzymes such as histone deacetylases (Supplementary Table 5).

The bidentate interaction between GSK-J1 and the catalytic metal is critical for enzyme inhibition. The pyridine regio-isomer GSK-J2 (Fig. 2a) cannot form such an interaction and hence showed considerably weaker JMJD3 inhibition ($IC_{50} > 100$ μM, Supplementary Table 3). Because GSK-J2 has analogous physicochemical properties to GSK-J1 but lacks H3K27 demethylase activity, it provides an ideal inactive control molecule for elucidating the functional role of JMJD3 inhibition.

Given the unusual binding mode of GSK-J1, we investigated its selectivity and its ability to engage full-length JMJD3 and UTX in an endogenous cellular environment by using a chemoproteomics approach. This approach combines a competition-binding assay in a cell extract, using an immobilized analogue of the inhibitor, with western blotting and mass spectrometry analysis[19]. The co-crystal structure of GSK-J1 with JMJD3 indicated that substitution at the *para* position to the pyridine nitrogen would enable access to solvent and would hence be amenable to immobilization, yielding GSK-J3 (Fig. 3c). GSK-J3 retained good activity against JMJD3 and was attached to sepharose beads to generate a KDM6 probe matrix. This
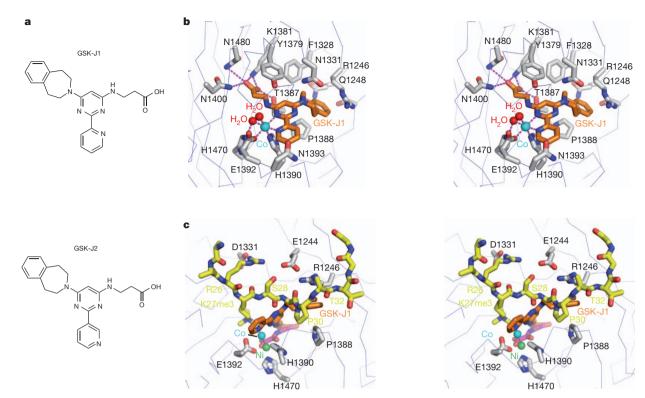
Figure 2 | Structure of the inhibitor GSK-J1 bound in the catalytic pocket of human JMJD3. a, Chemical structures of GSK-J1 and GSK-J2. **b**, Stereo view of the intermolecular interactions between the bound inhibitor GSK-J1 (orange) and the human JMJD3 enzyme residues (grey) that line the catalytic pocket, with hydrogen bonds shown as dashed pink lines. The heteroatoms in the stick representation are coloured according to atom type: oxygen (red) and nitrogen (blue). **c**, Stereo view of the superposition of the bound H3K27me3 peptide and the bound inhibitor GSK-J1 in their respective complexes with JMJD3. For clarity, the histone peptide (yellow), NOG (pink), Ni^{2+} (green) and certain JMJD3 side chains (grey) are shown only from the peptide complex, and GSK-J1 (orange) and Co^{2+} (blue) are shown from the inhibitor complex.

matrix precipitated Flag-tagged full-length JMJD3 and UTX from transiently transfected HEK-293 cells (Fig. 3d). This interaction with JMJD3 and UTX was specific, because the addition of free GSK-J1 to the lysate completely inhibited the binding of JMJD3 and UTX to the probe matrix (Fig. 3d). To confirm the ability of GSK-J1 to bind to endogenous JMJD3 in a relevant physiological context, we next used phorbol myristate acetate (PMA)-stimulated HL-60 monocytic cells to induce the expression of JMJD3, which was then specifically captured by the GSK-J3 probe matrix (Fig. 3e). Furthermore, in quantitative mass spectrometric experiments, JMJD3 was the only protein that was specifically captured by the probe matrix, as indicated by the inhibition of its binding in the presence of free GSK-J1 (Fig. 3f), thus confirming the remarkable selectivity of GSK-J1 in a cellular context.

Although critical for *in vitro* binding (Fig. 2b), the highly polar carboxylate group of GSK-J1 restricts cellular permeability. Therefore, we used a pro-drug strategy, masking the polarity of the acid groups of GSK-J1 and GSK-J2 with ethyl esters and yielding GSK-J4 and GSK-J5 (Fig. 4a). These cell-penetrating esters (with *in vitro* potency in the mass spectrometry assay, $IC_{50} > 50\,\mu M$) are rapidly hydrolysed by macrophage esterases, thereby generating pharmacologically relevant intracellular concentrations of GSK-J1 and GSK-J2 (Table 1). The ethyl ester pro-drug GSK-J4 was confirmed to have cellular activity in Flag–JMJD3-transfected HeLa cells, in which GSK-J4 prevented the JMJD3-induced loss of nuclear H3K27me3 immunostaining (Fig. 4b and Supplementary Fig. 20). Administration of GSK-J4 increased total nuclear H3K27me3 levels in untransfected cells (Supplementary Fig. 20c).

JMJD3 is involved in several physiological functions, including the inflammatory response[20–23]. In macrophages, JMJD3 expression is rapidly induced by proinflammatory stimuli through a nuclear factor-κB (NF-κB)-dependent mechanism, and JMJD3 is recruited to the transcription start sites (TSSs) of over 70% of lipopolysaccharide (LPS)-induced genes, where it participates directly in the transcriptional

response[24,25]. Crucially, it remains unclear whether this activation of transcription is achieved through the demethylation of H3K27me3 at target gene promoters[24]. Therefore, we next examined the efficacy of GSK-J4 and GSK-J5 at inhibiting the LPS-induced response of human primary macrophages derived from healthy volunteers. Administration of GSK-J4 significantly reduced the expression of 16 of 34 LPS-driven cytokines as assessed by PCR array (Fig. 4c and Supplementary Table 6), including tumour-necrosis factor-α (TNF-α).

TNF-α is important in a variety of inflammatory disorders, prompting us to make a more detailed characterization of GSK-J4's control of this cytokine. TNF-α protein production was inhibited in a dose-dependent manner (IC_{50} of GSK-J4 for TNF-α blockade, 9 μM), whereas the inactive isomer, GSK-J5, had no effect (Fig. 4d). No interference from cellular toxicity was observed under these conditions (Supplementary Fig. 21a); upstream NF-κB signalling seemed to be intact (Fig. 4e); and neither GSK-J4 nor GSK-J5 directly affected the levels of LPS-induced JMJD3 or of LPS-independent UTX and EZH2 (Supplementary Fig. 21b). Kinetic analysis suggested that the early phases of TNF-α induction were less affected by our inhibitor than were the later phases (Supplementary Fig. 21c, d). Interestingly, short interfering RNA (siRNA)-mediated knockdown of either JMJD3 or UTX alone in human primary macrophages had no effect on TNF-α production. Under these conditions, GSK-J4 was still able to inhibit TNF-α production, suggesting redundancy within the KDM6 subfamily in the regulation of this cytokine. To mimic the pan-KDM6 activity of GSK-J4, we performed dual depletion of JMJD3 and UTX in these cells, and this depletion eliminated TNF-α production (Fig. 4f and Supplementary Fig. 22). In addition, chromatin immunoprecipitation (ChIP) studies confirmed that GSK-J4, but not GSK-J5, prevented the LPS-induced loss of H3K27me3 associated with the *TNFA* TSS and blocked the recruitment of RNA polymerase II to this locus (Fig. 4g). Neither GSK-J4 nor GSK-J5 had an effect on H3K27 methylation
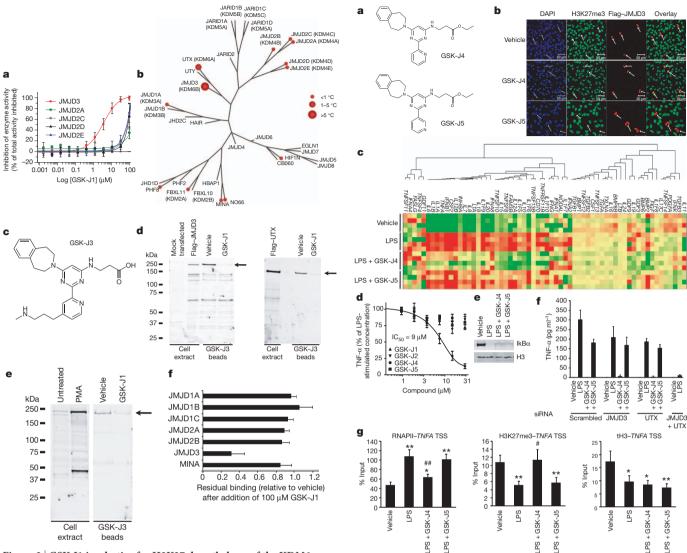
Figure 3 | GSK-J1 is selective for H3K27 demethylases of the KDM6 subfamily and specifically binds to endogenous JMJD3. a, Evaluation of the selectivity of GSK-J1 in JMJ mass spectrometric assays. Data are presented as the mean ± s.d. **b**, Phylogenetic tree of human JMJ enzymes illustrating the selectivity of GSK-J1 for demethylases of the KDM6 subfamily over other KDM subfamilies of methyl-lysine demethylases, as determined by melting point (T_m) shift screening. Temperature differences (°C) are shown in red circles. Only UTX and JMJD3 show significant stabilization (>1 °C), indicating ligand binding. **c–f**, Chemoproteomics studies show target engagement in a cellular environment. **c**, GSK-J3, an amine analogue of GSK-J1 used for bead immobilization. **d**, Immobilized GSK-J3 was able to affinity-capture Flag-tagged JMJD3 and Flag-tagged UTX from transiently transfected HEK-293 cells. Western blotting analysis with an anti-Flag antibody shows the capture of JMJD3 (left, arrow) and UTX (right, arrow), both of which were competitively inhibited by the addition of 100 μM GSK-J1. **e**, PMA treatment of HL-60 cells (80 nM for 16 h) induced the expression of JMJD3, as shown by western blotting. Immobilized GSK-J3 captured the endogenous JMJD3 protein (arrow), and this capture was competitively inhibited by 100 μM GSK-J1. **f**, GSK-J1 is selective for JMJD3, as shown by mass-spectrometry-based quantification of JMJC proteins (captured with GSK-J3 beads) from PMA-stimulated HL-60 cells in the presence of 100 μM GSK-J1. Data are presented as the mean from two experiments, and error bars represent the 95% confidence intervals.

Figure 4 | GSK-J1 inhibits TNF-α production by human primary macrophages in an H3K27-dependent manner. a, Chemical structures of the ethyl ester pro-drugs GSK-J4 and GSK-J5. **b**, Administration of 25 μM GSK-J4, but not GSK-J5, preserved nuclear H3K27me3 staining (green) in Flag–JMJD3-transfected (red, arrows) HeLa cells. Scale bars, 50 μm. **c**, Heat map representation of cytokine expression by human primary macrophages activated with LPS (for 2 h) in the presence of 30 μM GSK-J4 or GSK-J5 ($n = 4$ donors; green, low expression; red, high expression). **d**, TNF-α production by human primary macrophages activated with LPS in the presence of the indicated concentrations of GSK-J4 or GSK-J5 for 6 h (data are presented as the mean ± s.e.m. from $n = 5$ donors). **e**, Western blot showing the degradation of inhibitor of NF-κB α (IκBα) in whole cell lysates of LPS-stimulated human primary macrophages (30 min) in the presence of 30 μM GSK-J4 or GSK-J5. **f**, TNF-α production by LPS-stimulated human primary macrophages (1 h) transfected with scrambled, or JMJD3-directed or UTX-directed siRNA, in the presence of 30 μM GSK-J4 or GSK-J5 (data are presented as the mean ± s.d. from five transfection replicates from one representative experiment). **g**, ChIP analysis of the association of RNA polymerase II (RNAPII), H3K27me3 and total H3 (tH3) with the *TNFA* TSS in LPS-stimulated human primary macrophages (1 h) in the presence of 30 μM GSK-J4 or GSK-J5 (data are presented as the mean ± s.e.m. from $n = 8$ donors). P values were calculated using a two-tailed Student's t-test. *, $P < 0.05$ compared with vehicle; **, $P < 0.01$ compared with vehicle; #, $P < 0.05$ compared with LPS; ##, $P < 0.01$ compared with LPS.

upstream of the *TNFA* TSS (Supplementary Fig. 23). Finally, GSK-J4, but not GSK-J5, blocked the production of TNF-α by macrophages derived from patients with rheumatoid arthritis (Supplementary Fig. 24).

Our findings provide a strategy for the rational design and characterization of catalytic site inhibitors of JMJ enzymes. Specifically, we exploited active site plasticity and targeted key substrate-binding

features to identify the first series of selective H3K27 inhibitors. A chemoproteomics approach was used to assess the engagement of endogenous targets and to confirm the remarkable selectivity of the inhibitors within a relevant cellular context. By using these inhibitors, we provide clear insight into the biological relevance of the demethylase

Table 1 | Intracellular delivery of GSK-J1 and GSK-J2

Compound	[GSK-J1] in lysate (μM)	[GSK-J2] in lysate (μM)
Acid (GSK-J1 or GSK-J2)	1.6 ± 1.6	7.8 ± 1.6
Ethyl ester (GSK-J4 or GSK-J5)	11.8 ± 0.6	17.4 ± 1.5

GSK-J1 and GSK-J2 were detected in the lysates of human primary macrophages 1 h after the administration of 30 μM GSK-J1 or GSK-J2 or 30 μM of the pro-drugs GSK-J4 or GSK-J5 (mean ± s.d. from $n = 3$ donors).

function of the KDM6 JMJ enzymes. Furthermore, we show that H3K27 demethylation and specifically JMJD3 and UTX catalytic activity are critical determinants of proinflammatory gene activation in human primary macrophages. In reporting the first selective KDM6 inhibitor, we show the relevance and tractability of demethylase inhibition, as well as indicate that the KDM6 class of proteins can be targeted by epigenetic focused drug discovery, which may have broad therapeutic application.

METHODS SUMMARY

The cloning, expression, protein purification and crystallization of mouse JMJD3(1132–1641), engineered mouse JMJD3(1157–1641)–H3, human JMJD3 and their complexes are described in the Supplementary Information. GSK-J1–GSK-J5 were synthesized as pure compounds and fully characterized as outlined in the Supplementary Information. The fluorescent thermal shift assays, AlphaScreen, and RapidFire and matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometric assays to measure the inhibition of histone demethylase activity and JMJ selectivity, as well as the chemoproteomics experiments to establish compound–target engagement, were carried out as described in the Supplementary Information or as previously reported[19]. The procedures for the studies in human primary macrophages, including intracellular compound quantification, quantitative PCR with reverse transcription, enzyme-linked immunosorbent assays, western blotting, siRNA-mediated knockdown of JMJD3 and UTX, and ChIP are outlined in the Supplementary Information, together with a description of the antibodies and the materials.

1. Mosammaparast, N. & Shi, Y. Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem.* **79,** 155–179 (2010).
2. Klose, R. J. & Zhang, Y. Regulation of histone methylation by demethylimination and demethylation. *Nature Rev. Mol. Cell Biol.* **8,** 307–318 (2007).
3. Shi, Y. Histone lysine demethylases: emerging roles in development, physiology and disease. *Nature Rev. Genet.* **8,** 829–833 (2007).
4. Trojer, P. & Reinberg, D. Histone lysine demethylases and their impact on epigenetics. *Cell* **125,** 213–217 (2006).
5. Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biochem. Sci.* **36,** 7–18 (2011).
6. Shi, Y. & Whetstine, J. R. Dynamic regulation of histone lysine methylation by demethylases. *Mol. Cell* **25,** 1–14 (2007).
7. Pedersen, M. T. & Helin, K. Histone demethylases in development and disease. *Trends Cell Biol.* **20,** 662–671 (2010).
8. Kouzarides, T. Chromatin modifications and their function. *Cell* **128,** 693–705 (2007).
9. Yang, Y. *et al.* Structural insights into a dual-specificity histone demethylase ceKDM7A from *Caenorhabditis elegans. Cell Res.* **20,** 886–898 (2010).
10. Horton, J. R. *et al.* Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nature Struct. Mol. Biol.* **17,** 38–43 (2010).
11. Couture, J. F. *et al.* Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nature Struct. Mol. Biol.* **14,** 689–695 (2007).
12. Ng, S. S. *et al.* Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448,** 87–91 (2007).
13. Gehani, S. S. *et al.* Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Mol. Cell* **39,** 886–900 (2010).
14. Sengoku, T. & Yokoyama, S. Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. *Genes Dev.* **25,** 2266–2277 (2011).
15. Gamo, F. J. *et al.* Thousands of chemical starting points for antimalarial lead identification. *Nature* **465,** 305–310 (2010).
16. Yu M.. *et al.* Bisubstrate inhibition: theory and application to *N*-acetyltransferases. *Biochemistry* **45,** 14788–14794 (2006).
17. Mackeen, M. M. *et al.* Small-molecule-based inhibition of histone demethylation in cells assessed by quantitative mass spectrometry. *J. Proteome Res.* **9,** 4082–4092 (2010).
18. Sakurai, M. *et al.* A miniaturized screen for inhibitors of Jumonji histone demethylases. *Mol. Biosyst.* **6,** 357–364 (2010).
19. Bantscheff, M. *et al.* Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nature Biotechnol.* **29,** 255–265 (2011).
20. Lan, F. *et al.* A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449,** 689–694 (2007).
21. Lee, M. G. *et al.* Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318,** 447–450 (2007).
22. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449,** 731–734 (2007).
23. Miller, S. A., Mohn, S. E. & Weinmann, A. S. Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol. Cell* **40,** 594–605 (2010).
24. De Santa, F. *et al.* Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. *EMBO J.* **28,** 3341–3352 (2009).
25. De Santa, F. *et al.* The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* **130,** 1083–1094 (2007).

Author Contributions L.K., C.-w.C., J.L., Z.C., G.D., U.O., D.J.P., K.C., K.L. and D.M.W. defined the research, interpreted the data and wrote the manuscript. All authors performed the experiments, analysed the data and/or contributed to the scientific discussion.

LETTER

The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming

Abed AlFatah Mansour[1]*, Ohad Gafni[1]*, Leehee Weinberger[1], Asaf Zviran[1], Muneef Ayyash[2], Yoach Rais[1], Vladislav Krupalnik[1], Mirie Zerbib[1], Daniela Amann-Zalcenstein[3,4], Itay Maza[1], Shay Geula[1], Sergey Viukov[1], Liad Holtzman[1], Ariel Pribluda[1], Eli Canaani[5], Shirley Horn-Saban[4], Ido Amit[3], Noa Novershtern[1]* & Jacob H. Hanna[1]

Induced pluripotent stem cells (iPSCs) can be derived from somatic cells by ectopic expression of different transcription factors, classically Oct4 (also known as Pou5f1), Sox2, Klf4 and Myc (abbreviated as OSKM)[1]. This process is accompanied by genome-wide epigenetic changes[2–5], but how these chromatin modifications are biochemically determined requires further investigation. Here we show in mice and humans that the histone H3 methylated Lys 27 (H3K27) demethylase Utx[6–9] (also known as Kdm6a) regulates the efficient induction, rather than maintenance, of pluripotency. Murine embryonic stem cells lacking Utx can execute lineage commitment and contribute to adult chimaeric animals; however, somatic cells lacking Utx fail to robustly reprogram back to the ground state of pluripotency. Utx directly partners with OSK reprogramming factors and uses its histone demethylase catalytic activity to facilitate iPSC formation. Genomic analysis indicates that Utx depletion results in aberrant dynamics of H3K27me3 repressive chromatin demethylation in somatic cells undergoing reprogramming. The latter directly hampers the derepression of potent pluripotency promoting gene modules (including Sall1, Sall4 and Utf1), which can cooperatively substitute for exogenous OSK supplementation in iPSC formation. Remarkably, Utx safeguards the timely execution of H3K27me3 demethylation observed in embryonic day 10.5–11 primordial germ cells (PGCs)[10], and Utx-deficient PGCs show cell-autonomous aberrant epigenetic reprogramming dynamics during their embryonic maturation in vivo. Subsequently, this disrupts PGC development by embryonic day 12.5, and leads to diminished germline transmission in mouse chimaeras generated from Utx-knockout pluripotent cells. Thus, we identify Utx as a novel mediator with distinct functions during the re-establishment of pluripotency and germ cell development. Furthermore, our findings highlight the principle that molecular regulators mediating loss of repressive chromatin during in vivo germ cell reprogramming can be co-opted during in vitro reprogramming towards ground state pluripotency.

We established primed epiblast stem cells (EpiSCs)[11,12] carrying a Nanog–GFP reporter and doxycycline-inducible OSKM alleles, and applied piggyBac (PB)-based gene-trap screening[13] to identify inhibitors of EpiSC reversion to naive pluripotency (Supplementary Fig. 1a–c). Several clonal populations analysed carried a disruptive PB integration in the second intron of the Utx gene and failed to generate Nanog–GFP+ naive pluripotent cells adequately (Fig. 1a and Supplementary Fig. 1d). Functional rescue by overexpression of Utx induced rapid Nanog–GFP reactivation (Fig. 1a).

Utx is an X-chromosome-encoded gene that, together with Uty and Jmjd3 (also known as Kdm6b), belongs to a small family of JmjC-domain containing enzymes that mediate demethylation of H3K27 tri- and dimethyl (me3/2) repressive chromatin marks[6–9].
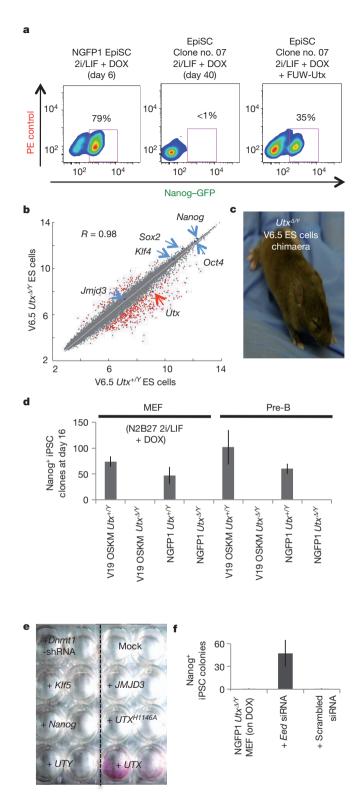
We knocked-out Utx in V6.5 mouse embryonic stem (ES) cell line and examined their pluripotent characteristics (Supplementary Fig. 2). V6.5 $Utx^{\Delta/Y}$ ES cells were pluripotent, as evidenced by normal staining for pluripotency markers, global gene expression, and growth rate and lineage commitment marker upregulation upon differentiation (Fig. 1b and Supplementary Figs 3, 4). Whereas Utx-deficient embryos predominantly die around embryonic day (E)10.5 (ref. 14), and at E8.5–E9.5 following tetraploid complementation assay (Supplementary Fig. 5), viable high-contribution adult chimaeras were readily obtained from $Utx^{\Delta/Y}$ ES cells (Fig. 1c and Supplementary Fig. 4). These results demonstrate that Utx is dispensable for pluripotent state maintenance, and that functional commitment into multiple somatic lineages can be achieved in the absence of this gene.

We next tested whether Utx may be required for the re-establishment, rather than maintenance, of pluripotency. Thus, we knocked out Utx in two different 'secondary reprogramming' pluripotent stem cell lines (NGFP1 iPSCs and V19 ES cells) carrying doxycycline-inducible OSKM transgenes[15]. Genetically matched NGFP1 or V19 $Utx^{\Delta/Y}$ and $Utx^{+/Y}$ cells were injected into host blastocysts, and differentiated transgenic mouse embryonic somatic fibroblasts (MEFs) and pre-B cells were isolated. Induction of pluripotency in these cells by doxycycline demonstrated a marked block in retrieving iPSCs from $Utx^{\Delta/Y}$ cells in naive and primed reprogramming conditions (Fig. 1d and Supplementary Fig. 6). Long-term follow-up[15,16] of reprogramming of monoclonal populations of donor somatic $Utx^{\Delta/Y}$ B cells showed a severely delayed detection of Nanog–GFP reactivation following doxycycline induction (Supplementary Fig. 7). Notably, reprogrammed Nanog–GFP+ cells that could be derived at low efficiency and delayed kinetics retained aberrant global H3K27me3 and H3K4me3 epigenetic profiles (Supplementary Fig. 8), consistent with aberrant reprogramming. Human fibroblasts demonstrated reduced iPSC formation upon UTX knock-down (Supplementary Fig. 9). Furthermore, the lack of Utx inhibited the epigenetic reversion of primed EpiSCs and reprogramming via cell fusion (Supplementary Fig. 10). These results demonstrate that the absence of Utx demethylase constitutes a major roadblock for re-establishing pluripotency in vitro.

Utx is ubiquitously expressed in somatic tissues (Supplementary Fig. 5b), and its ectopic overexpression in addition to OSKM does not lead to increased efficiency in iPSC formation (Supplementary Fig. 11a, b). Acute tamoxifen-induced ablation of Utx in NGFP1-CreER; $Utx^{flox/Y}$ transgenic reprogrammable cells at different partially reprogrammed stages[17] resulted in diminished progression towards pluripotency (Supplementary Fig. 11c). However, once the cells had reactivated the endogenous Nanog–GFP reporter (characterized to constitute the final molecular barrier to pluripotency[18]), deleting Utx with tamoxifen had no influence on obtaining iPSCs in vitro (Supplementary Fig. 11c). These results indicate that Utx expression

[1]The Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel. [2]The Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel. [3]The Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel. [4]The Department of Biological Services—Genomics Unit, Weizmann Institute of Science, Rehovot 76100, Israel. [5]The Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel.
*These authors contributed equally to this work.

a

NGFP1 EpiSC
2i/LIF + DOX
(day 6)

EpiSC
Clone no. 07
2i/LIF + DOX
(day 40)

EpiSC
Clone no. 07
2i/LIF + DOX
+ FUW-Utx

79%

<1%

35%

PE control

Nanog–GFP

Figure 1 | Utx promotes the re-establishment of pluripotency in somatic cells. a, Fluorescence-activated cell sorting (FACS) analysis for Nanog–GFP reporter conducted on the indicated EpiSC-derived cell samples and mutant clones grown in 2i/LIF (2i/LIF is media containing LIF and two small-molecule inhibitors: PD0325901 (ERK1/2 inhibitor) and CHIR99021 (GSK-3β inhibitor)) plus doxycycline (DOX) conditions. PE, phycoerythrin. **b**, Scatter plot of transcriptional levels measured in $Utx^{+/Y}$ (x-axis, $n = 1$) and $Utx^{\Delta/Y}$ (y-axis, $n = 1$) ES cells. Transcription levels are shown in $\log_2$ scale. Rank correlation between the samples is shown. **c**, Adult chimaeria as indicated by the agouti coat colour. **d**, Reprogramming of 5,000 secondary OSKM-transgenic MEF or pre-B donor cells after 16 days on doxycycline in naive growth conditions, and quantifying Nanog$^+$ iPSC colonies ($n = 3$). N2B27, serum-free media (see Methods). **e**, $Utx^{\Delta/Y}$ somatic cells on doxycycline for 7 days were transduced with different rescue overexpression vectors and analysed 7 days later after passage for iPSC formation via staining for alkaline phosphatase. **f**, As in **e**, except rescue was conducted by Eed siRNA transfection ($n = 2$). Error bars indicate s.d.

b

$R = 0.98$

Nanog
Sox2
Klf4
Oct4
Jmjd3
Utx

V6.5 $Utx^{\Delta/Y}$ ES cells

V6.5 $Utx^{+/Y}$ ES cells

c

$Utx^{\Delta/Y}$
V6.5 ES cells
chimaera

d

MEF   Pre-B

(N2B27 2i/LIF
+ DOX)

Nanog$^+$ iPSC clones at day 16

V19 OSKM $Utx^{+/Y}$
V19 OSKM $Utx^{\Delta/Y}$
NGFP1 $Utx^{+/Y}$
NGFP1 $Utx^{\Delta/Y}$
V19 OSKM $Utx^{+/Y}$
V19 OSKM $Utx^{\Delta/Y}$
NGFP1 $Utx^{+/Y}$
NGFP1 $Utx^{\Delta/Y}$

e

+Dnmt1 -shRNA | Mock
+ Klf5 | + JMJD3
+ Nanog | + UTX^{H1146A}
+ UTY | + UTX

f

Nanog$^+$ iPSC colonies

NGFP1 $Utx^{\Delta/Y}$ MEF (on DOX)
+ Eed siRNA
+ Scrambled siRNA

(Fig. 1e)[7]. We next tested whether inducing global loss of H3K27me1/2/3 could facilitate the reprogramming of $Utx^{\Delta/Y}$ somatic cells. Thus, we inhibited Eed expression, a component of the polycomb group protein repressive complex (PRC2) required for the establishment and maintenance of H3K27me marks, and its inhibition leads to genome-wide loss of H3K27me1/2/3 (refs 16, 19). Short interfering RNA (siRNA) for Eed resulted in the rapid derivation of iPSCs from $Utx^{\Delta/Y}$ somatic cells undergoing reprogramming following doxycycline induction (Fig. 1f). This indicates that Utx predominantly contributes to reprogramming towards naive pluripotency through its H3K27 demethylation activity, and that authentic reprogramming towards the ground state of pluripotency correlates with global loss of H3K27me3 deposition. However, our results do not exclude synergistic contributions from Jmjd3, or demethylation-independent functions of Utx in reprogramming.

We next conducted gene expression profiling of somatic cells during reprogramming and compared them to patterns observed in ES cells. The effect of Utx deletion on the induction, rather than on the maintenance, of the pluripotent state was evident at the transcriptional level. The number of genes affected by Utx knockout was 6.6 times higher in MEF samples after doxycycline induction, compared to ES cells (Supplementary Fig. 13a and Supplementary Table 1). Hierarchical clustering and principle component analysis clustered $Utx^{\Delta/Y}$ doxycycline-induced MEFs together with untreated $Utx^{\Delta/Y}$ and $Utx^{+/Y}$ MEFs (Fig. 2a and Supplementary Fig. 13b), thus indicating that $Utx^{\Delta/Y}$ MEFs fail to initiate molecular changes typically detected after 8 days of OSKM induction, and remain highly similar to the original donor somatic cells (Pearson correlation = 0.97). Importantly, the genes that are activated (>2-fold) during reprogramming in $Utx^{+/Y}$ MEFs are enriched (false discovery rate (FDR) <1%) for targets of Klf4, Oct4 and Sox2 factors, as was measured in ES cells or during iPSC formation after OSKM induction in MEFs[20] (Fig. 2b). In contrast, known targets of the latter reprogramming factors were preferentially enriched among the downregulated gene subgroup in $Utx^{\Delta/Y}$ doxycycline-induced MEFs (>2-fold) (Fig. 2b). In both cases, the affected genes were highly enriched for H3K27me3 chromatin marks that were aberrantly retained in $Utx^{\Delta/Y}$ MEFs following doxycycline induction (Supplementary Table 2).

We next mapped H3K27me3 and H3K4me3 chromatin modifications by genome-wide chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) analysis in MEFs before and after doxycycline induction in genetically matched $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ samples and in established ES cells (Supplementary Table 3). The perturbed active H3K4me3 and repressive H3K27me3 chromatin marks correlated with global active and repressed transcriptional levels, respectively[2] (Supplementary Fig. 13c). The effect of Utx deletion during iPSC reprogramming was also apparent at the global chromatin level. Twenty-seven per cent of the 14,619 mapped genes had aberrant H3K27me3 and/or H3K4me3 modifications in $Utx^{\Delta/Y}$ cells (Fig. 2c). A total of 957 genes retained or aberrantly gained H3K27me3 repressive marks (Fig. 2c) during reprogramming in

is not a rate-limiting factor for iPSC formation, but is rather essential and continuously required during *in vitro* reprogramming. Specific knockdown of Jmjd3 did not significantly reduce iPSC formation (Supplementary Figs 9, 12). Furthermore, Utx-deficient somatic cell reprogramming could not be reconstituted by overexpression of JMJD3, UTY or other tested pluripotency promoting factors, suggesting a unique requirement for Utx in this process (Fig. 1e). Importantly, $Utx^{\Delta/Y}$ iPSC formation rescue was observed by overexpressing wild-type UTX, but not the UTX^{H1146A} allele, which carries a mutation in the JmjC catalytic domain, abrogating its H3K27 demethylase activity
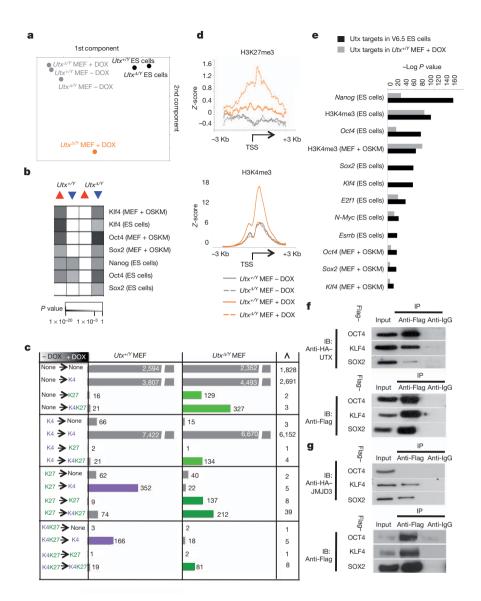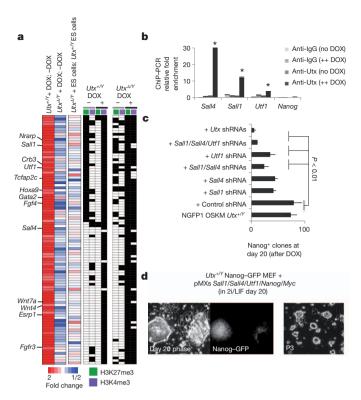
Figure 2 | Genomic dissection of reprogramming in Utx-knockout cells. a, Principle component analysis of transcriptional profiles of the indicated samples. Colours indicate distinct clusters: ES cells (black); MEF and MEF-like (grey); and doxycycline (DOX)-induced partially reprogrammed pre-iPSCs (orange). **b**, Selected functional categories enriched in induced (red) and repressed (blue) genes (>2-fold), following 8 days of doxycycline in $Utx^{+/Y}$ and $Utx^{Δ/Y}$ MEFs. Grey intensity indicates P value. **c**, Number of genes with H3K4me3 or H3K27me3 before and after doxycycline induction in $Utx^{+/Y}$ and $Utx^{Δ/Y}$ MEFs. Light green indicates number of genes that gained H3K27me3 in $Utx^{Δ/Y}$ only; dark green indicates number of gene that retained H3K27me3 in $Utx^{Δ/Y}$ only; purple indicates number of gene that gained/maintained H3K4me3 in $Utx^{+/Y}$ only. Right column shows overlapping gene number in each category. **d**, Binding profile around the transcription start site (TSS) of H3K4me3 and H3K27me3. Profiles were calculated over 805 genes that are bound by at least 3 of the following factors: Oct4, Sox2, Klf4, Myc and Nanog. Z-score of the mean profile is plotted (y-axis) against a sliding window around the TSS (x-axis). **e**, Top categories enriched among genes bound by Utx in ES cells (black) or in MEFs + DOX (grey), following genome-wide ChIP-seq analysis conducted either on ES cells or MEFs after OSKM transduction (as indicated in brackets). −Log P value is presented for each category. **f, g**, Co-imuunoprecipoitation (IP) of Flag-tagged OCT4, KLF4 or SOX2 with HA-tagged UTX or JMJD3, followed by immunoblot analysis (IB).

$Utx^{Δ/Y}$ and not in $Utx^{+/Y}$ MEFs. This suggests that Utx also safeguards against aberrant acquisition of repressive H3K27me3 marks after OSKM induction[2]. More than 500 genes became potentially active by losing H3K27me3 and gaining or maintaining H3K4me3 active chromatin marks in $Utx^{+/Y}$, but not in $Utx^{Δ/Y}$ doxycycline-induced MEFs (Fig. 2c). Notably, Utx deletion resulted in global increase in H3K27me3 in doxycycline-induced MEFs but not in ES cells (Supplementary Fig. 14a). This trend is consistent with our functional data indicating that Utx predominantly regulates pluripotency re-establishment rather than maintenance. H3K4me3 was decreased in doxycycline-induced MEFs but not in ES cells (Supplementary Fig. 14a), possibly resulting from an indirect effect of perturbation in H3K27me3-repressed genes and/or from the fact that Utx partners with the Mll2/3–H3K4me3-transferase complex[6,7]. Consistent with the gene-expression enrichment analysis in Fig. 2b, known OSKM and Nanog directly bound target genes in ES cells and MEFs undergoing reprogramming[20] were found to be methylated for H3K27me3 and demethylated for H3K4me3 in $Utx^{Δ/Y}$ doxycycline-induced MEFs (Fig. 2d). Overall, the above analysis provides evidence for aberrant transcriptional and epigenetic reprogramming dynamics upon OSKM induction in Utx-knockout somatic cells, and suggests that genes regulated by the pluripotency reprogramming machinery cannot be adequately reactivated in the absence of Utx.

ChIP-seq analysis for Utx binding in ES cells identified 2,416 bound genes that are enriched for H3K4me3 active chromatin marked genes as well as for targets of Oct4, Sox2 and Klf4 (Fig. 2e). Motif search analysis in Utx-binding locations identified predicted binding motifs for Klf4, Sox2, Oct4 and Nanog as the four most highly enriched (Supplementary Fig. 14b). ChIP-seq analysis for Utx binding in doxycycline-induced MEFs showed that Utx specifically bound 1,845 target genes that are enriched for H3K4me3 active chromatin marked genes as well as for targets of Nanog and Oct4 (Fig. 2e). Utx targets in MEFs following doxycycline induction, aberrantly accumulated H3K27me3 marks in these loci in $Utx^{Δ/Y}$ cells (Supplementary Fig. 14c). The above findings prompted us to examine whether Utx physically associates with the OSK reprogramming proteins. We established that Flag-tagged OCT4, KLF4 and SOX2 specifically co-immunoprecipitated with haemagglutin (HA)-tagged UTX (Fig. 2f). JMJD3 was not able to interact directly with OCT4, partially explaining the distinct influence for different H3K27 demethylases on iPSC formation (Fig. 2g).

We then aimed to delineate unbiasly some of the key genes that are deregulated early during the reprogramming process by the loss of Utx. We identified 1,430 genes that were transcriptionally upregulated (>2-fold) in $Utx^{+/Y}$ and not in $Utx^{Δ/Y}$ MEFs undergoing reprogramming. Of these, 98 genes had aberrantly upregulated H3K27me3

Figure 3 | Downstream effectors of Utx that promote iPSC reprogramming. a, Unbiased identification of 98 deregulated genes that were induced (>2-fold) in $Utx^{+/Y}$ but not in $Utx^{\Delta/Y}$. Left, expression change before and after 8 days of doxycycline (DOX) induction in $Utx^{+/Y}$ and $Utx^{\Delta/Y}$ MEFs, and in ES cells (comparing $Utx^{\Delta/Y}$ to $Utx^{+/Y}$). Right, chromatin marks before and after doxycycline induction. Purple indicates H3K4me3 modifications; green indicates H3K27me3 modifications; black boxes, methylated; white boxes, unmethylated. **b**, Anti-Utx and control anti-IgG ChIP–PCR relative fold enrichment of *Sall4*, *Sall1*, *Utf1* and *Nanog* loci in $Utx^{+/Y}$ and $Utx^{+/Y}$ MEFs before and after doxycycline treatment (normalized to $Utx^{+/Y}$ without doxycycline) ($n = 3$). **c**, Reprogramming following individual and combined knockdown with *Sall4*, *Sall1* and *Utf1* were introduced in $Utx^{+/Y}$ Nanog–GFP secondary MEFs ($n = 2$). *P value < 0.01 in comparison to wild-type samples. Eror bars indicate s.d. **d**, Overexpression of *Sall4*, *Sall1*, *Utf1*, *Nanog* and *Myc* reproducibly gives rise to Nanog–GFP+ iPSCs in $Utx^{+/Y}$ cells without exogenous OSK. P3, passage 3.

wild-type cells[2]. These genes were directly bound by Utx in ES cells and in doxycycline-induced MEFs, and aberrantly retained H3K27me3 in $Utx^{\Delta/Y}$ cells, as was confirmed by ChIP-seq and ChIP–polymerase chain reaction (PCR) analyses (Fig. 3b and Supplementary Figs 14d, 15–16a). This suggested that Utx can act during early reprogramming by directly regulating the reactivation of a variety of pluripotency promoting factors including Sall1, Sall4 and Utf1. Indeed, inhibiting the endogenous expression of *Sall1*, *Sall4* and *Utf1* in $Utx^{+/Y}$ MEFs through specific knockdown during reprogramming, recapitulates the block in iPSC formation observed in $Utx^{\Delta/Y}$ cells (Fig. 3c). Overexpression of Sall1, Sall4 and Utf1 in $Utx^{\Delta/Y}$ MEFs promoted the activation of early reprogramming marker SSEA1 (Supplementary Fig. 16b), but did not result in reactivation of Nanog–GFP, suggesting that other Utx-regulated pluripotency factors need to be reactivated to complete the reprogramming progression. This is consistent with the role of Utx as global regulator that targets multiple genes (Fig. 2e). Remarkably, consistent with a profound pluripotency promoting potential for Sall1, Sall4 and Utf1, their combined overexpression with Nanog and Myc in wild-type $Utx^{+/Y}$ MEFs replaces exogenous OSK expression, and generated chimaera-formation-competent iPSC clones (Fig. 3d). Overall, these data
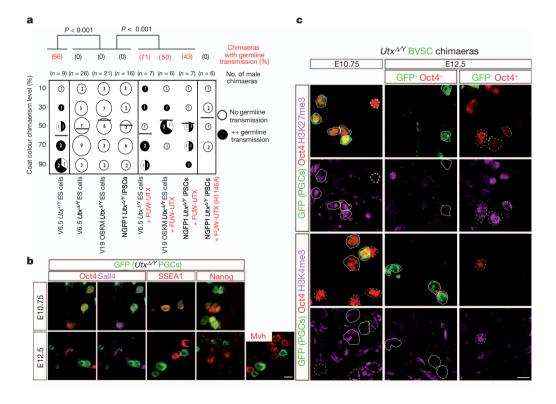
marks in $Utx^{\Delta/Y}$ doxycycline-induced MEFs (Fig. 3a), suggesting that their deregulation is likely to be caused by H3K27me3 retention. The signature of these deregulated genes is enriched (FDR <1%) for stem-cell maintenance factors (for example, *Fgf4*, *Sall4*, *Sall1* and *Utf1*), for validated direct targets of Klf4 and Oct4 (for example, *Sall1*, *Crb3* and *Nef1*), and for H3K27me3-marked genes (Supplementary Table 4). Of these, we focused on three transcription factors, *Sall1l*, *Sall4* and *Utf1*, previously shown to undergo reactivation after OSKM induction in

Figure 4 | Utx regulates epigenetic reprogramming and development of PGCs. a, Quantification of coat colour chimaerism in male mice derived from the indicated cell lines. Each pie chart indicates a group with the indicated chimaerism levels. Black areas indicate successful germline transmission, white areas indicate germline transmission failure. P values are indicated. **b, c**, Confocal images represent gonad sections from E10.75 or E12.5 chimaeric embryos injected with $Utx^{\Delta/Y}$ BVSC ES cells and immunostained for the indicated proteins. **b**, Immunostaining for GFP (labelling BVSC $Utx^{\Delta/Y}$ injected cells) and the indicated PGC markers revealed abnormal *in vivo* differentiation of *Utx*-deficient PGCs at E12.5 but not at E10.75. **c**, Immunofluorescence staining for epigenetic markers on the genital ridge of E10.75 embryos demonstrates that high levels of H3K27me3 are retained in *Utx*-deficient PGCs (unbroken circles) compared with adjacent wild-type PGCs (dashed circles). At E12.5 *Utx*-deficient PGCs have diminished levels of all tested epigenetic marks. Scale bars, 10 μm.

indicate that Utx can associate with OKS by a protein–protein interaction and promotes the reactivation of potent pluripotency promoting modules that cooperatively facilitate iPSC formation.

We next investigated whether Utx has a role in epigenetic reprogramming during early development. Whereas ES cells could be readily derived from $Utx^{Δ/Y}$ blastocysts (Supplementary Fig. 17), we noticed that chimaeric animals generated from three different $Utx^{Δ/Y}$ iPSC and ES cell clones reproducibly failed to demonstrate germline transmission after extensive breeding (Fig. 4a). Importantly, multiple adult chimaeras generated from three different corrected transgenic lines with catalytically active UTX, but not the mutant UTX^{H1146A} allele, robustly achieved germline transmission (Fig. 4a). This indicates that the defect in germ-lineage development can be unambiguously assigned to the absence of catalytically active Utx.

PGCs increase global levels of H3K27me3 by E8.5–E9.5, and then undergo a second wave of epigenetic reprogramming at E10.5–11 that entails global loss of the H3K27me3 mark[10]. Thus we tested whether Utx deficiency results in perturbed PGC development and reprogramming *in vivo* during this critical developmental window. To reproducibly study PGC development beyond E9.5–E10.5, we analysed chimaeric animals after microinjection of either V6.5 $Utx^{Δ/Y}$ ES cells labelled with LacZ, or the BVSC ES cell line[21] knocked out for Utx ($Utx^{Δ/Y}$; BVSC), which carries Blimp1–mVenus (Blimp1 also known as Prdm1) and Stella–CFP (Stella also known as Dppa3) early PGC reporters. $Utx^{Δ/Y}$ PGCs were found to express the PGC markers Oct4, Nanog, Sall4 and SSEA1 normally at E10.75 (Fig. 4b). However, only Utx-null PGCs did not undergo H3K27me3 demethylation during this developmental window (Fig. 4c and Supplementary Fig. 18). The remaining chromatin marks tested were not different at this stage, suggesting that aberrantly accumulated H3K27me3 at this critical developmental stage may be an early identifiable trigger for subsequent PGC developmental defects. Indeed, despite the adequate localization of Blimp1- and Stella-positive $Utx^{Δ/Y}$ PGCs in the gonads at E12.5, they fail to retain the expression of several PGC pluripotency markers (Nanog, Sall4, Oct4 and SSEA1) or show reactivation of the late PGC marker Mvh (Fig. 4b and Supplementary Figs 18, 19). In addition, $Utx^{Δ/Y}$ PGCs show aberrant chromatin dynamics at E12.5, as evident from the low chromatin levels of H3K9me3, H3K4me2/3 and H3K27me3, and reduced nuclear size in comparison to adjacent wild-type PGCs (Fig. 4c and Supplementary Fig. 18c). Finally, we find that $Utx^{Δ/Y}$ PGCs isolated from E8.5 chimaeras are deficient in their ability to revert to ground state pluripotency and form ES-like embryonic germ cells (Supplementary Figs 20, 21). Taken together, these results provide direct evidence for cell-autonomous aberrant germ cell development *in vivo* and perturbed epigenetic reprogramming in the absence of Utx.

In summary, we identify Utx as a critical regulator acting at molecular switches during reprogramming to ground state pluripotency to safeguard an efficient, timely and authentic execution of H3K27 demethylation that is critical for this process. Future experimental work is needed to contrast the developmental potency phenotypes in mice and pluripotent cells deficient for different combinations of components regulating the deposition of H3K27 and H3K4 methylation (for example, Utx, Jmjd3, Uty and Mll2/3 enzymes).

METHODS SUMMARY

The following knockout stem cell lines for Utx were generated by targeted disruption of the endogenous Utx locus via homologous recombination: (1) V6.5 ES cells; (2) NGFP1 iPSCs; (3) V19-OSKM ES cells; and (4) BVSC ES cells[21]. Full details of stem cell lines, cell culture, plasmids, siRNAs and antibodies used, as well as descriptions of methods for epigenetic reprogramming, immunofluorescence, immunoblotting, immunoprecipitation, embryo micromanipulation, PGC differentiation, bio-informatics and statistical analyses are provided in Methods.

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
2. Koche, R. P. *et al.* Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* **8**, 96–105 (2011).
3. Hemberger, M., Dean, W. & Reik, W. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature Rev. Mol. Cell Biol.* **10**, 526–537 (2009).
4. Singhal, N. *et al.* Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* **141**, 943–955 (2010).
5. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
6. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in *HOX* gene regulation and development. *Nature* **449**, 731–734 (2007).
7. Hong, S. *et al.* Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl Acad. Sci. USA* **104**, 18439–18444 (2007).
8. Sen, G. L., Webster, D. E., Barragan, D. I., Chang, H. Y. & Khavari, P. A. Control of differentiation in a self-renewing mammalian tissue by the histone demethylase JMJD3. *Genes Dev.* **22**, 1865–1870 (2008).
9. Lan, F. *et al.* A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689–694 (2007).
10. Hajkova, P. *et al.* Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877–881 (2008).
11. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
12. Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
13. Guo, G., Huang, Y., Humphreys, P., Wang, X. & Smith, A. A PiggyBac-based recessive screening method to identify pluripotency regulators. *PLoS ONE* **6**, e18189 (2011).
14. Lee, S., Lee, J. W. & Lee, S. K. UTX, a histone H3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. *Dev. Cell* **22**, 25–37 (2011).
15. Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595–601 (2009).
16. Montgomery, N. D. *et al.* The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr. Biol.* **15**, 942–947 (2005).
17. Stadtfeld, M., Maherali, N., Breault, D. T. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* **2**, 230–240 (2008).
18. Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722–737 (2009).
19. Hansen, K. H. *et al.* A model for transmission of the H3K27me3 epigenetic mark. *Nature Cell Biol.* **10**, 1291–1300 (2008).
20. Sridharan, R. *et al.* Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136**, 364–377 (2009).
21. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).

METHODS

Stem cell lines and cell culture. Reprogramming and maintenance of murine pluripotent cells were conducted in serum-free chemically defined N2B27-based media: 240 ml DMEM/F12 (Biological Industries, custom made), 240 ml Neurobasal (Invitrogen; 21103), 5 ml N2 supplement (Invitrogen; 17502048), 5 ml B27 supplement (Invitrogen; 17504044), 1 mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma), penicillin-streptomycin (Invitrogen), 5 mg ml^{-1} BSA (Sigma). Naive conditions for murine PSCs included 5 μg recombinant human LIF (Millipore; LIF1005) and small-molecule inhibitors CHIR99021 (3 μM; Axon Medchem) and PD0325901 (1 μM; TOCRIS). Primed N2B27 media for murine cells (EpiSCs) contained 8 ng ml^{-1} recombinant human bFGF (Peprotech Asia) and 20 ng ml^{-1} recombinant human activin (Peprotech Asia). Stem cell lines and mice deficient for Utx were generated by targeted disruption of the endogenous Utx locus via homologous recombination. The targeting strategy and construct (European Conditional Mouse Mutagenesis Program) introduced loxP sites spanning the third exon that would result in an out-of-frame and truncated product upon deletion, and introduced a LacZ reporter cassette driven by the endogenous Utx promoter. Fifty micrograms DNA of the targeting construct were linearized and electroporated into four different male pluripotent cell lines that were then subjected to selection with G418 (300 μg ml^{-1}) and gancyclovir: (1) V6.5 ES cells; (2) NGFP1 iPSCs[15]; (3) V19 OSKM ES cells; and (4) BVSC[21] ES cells. After 10 days of selection, double resistant clones were analysed for correct targeting of 5′ and 3′ arms (Supplementary Table 5). $Utx^{flox/Y}$ ES cells showed slightly reduced levels of Utx mRNA and protein, indicating that Utx^{flox} is a hypomorphic allele, but not a full null allele. Generation of the $Utx^{Δ}$ null allele was achieved after transfection with pPac-Cre construct, PSC picking and sub-cloning, followed by genomic DNA genotyping for detection of allele deletion by PCR (Supplementary Table 5). $Utx^{+/flox}$ females were crossed with PGK-Cre^m transgenic males to generate $Utx^{+/Δ}$; PGK-Cre$^+$ adult females or $Utx^{Δ/Y}$ embryos. $Utx^{+/Δ}$; PGK-Cre$^+$ females were crossed with $Utx^{flox/Y}$ males to obtain all expected genotypes.

Epigenetic reversion of primed epiblast cells and piggyBac mutagenesis screen. Male naive NGFP1 iPSCs[15] maintained in 2i/LIF conditions were injected into BDF2 blastocysts. Epiblasts from chimaeric embryos were dissected at E6.5 and explanted on fibronectin/vitronectin-coated plates in N2B27 bFGF/activin conditions supplemented with 1 μg ml^{-1} puromycin, allowing the isolation of NGFP1 EpiSCs. For epigenetic reversion of murine EpiSCs to naive pluripotency, cells were passaged into N2B27 2i/LIF conditions on vitronectin- (1 μg ml^{-1}) and gelatine- (0.2%) coated plates with or without doxycycline as indicated. When the epigenetic reversion assay involved single-cell plating, EpiSC growth medium was supplemented with ROCK inhibitor for 24 h before trypsinization. The piggyBac gene trap system was composed of gene trap vector, pGG85, and transposase expressing helper plasmid, pCAGG-PBase[13]. NGFP1 EpiSCs were grown in N2B27 bFGF/activin and ROCK inhibitor (10 μM) for 24 h and subsequently 5×10^6 cells were co-electroporated with both plasmids at 1:3 ratio (1 mg of pGG85 and 3 mg of helper plasmid) (Genepulser; Biorad) and allowed to recover in N2B27 bFGF/activin/ROCK inhibitor for 24 h. After which the cells were passaged in N2B27 2i/LIF plus G418 (150 μg ml^{-1}) plus doxycycline. At day 7 (bulk sort) and 14 (single-cell sort) cells were sorted for isolation and enriching for the Nanog–GFP-negative fraction and cells were maintained in N2B27 2i/LIF plus doxycycline. Genomic DNA from selected clones was subjected to Splinkerette PCR-based insertional analysis as previously described[13].

Reprogramming of somatic cells and cell transfection. Virus-containing supernatants of the different reprogramming viruses (FUW-tetO-lox-hKLF4, FUW-tetO-lox-hOCT4, FUW-tetO-lox-hSOX2, FUW-tetO-mKlf4, FUW-tetO-mOct4, FUW-tetO-mSox2, FUW-tetO-c-Myc) were supplemented with the FUW-lox-M2rtTA virus (when necessary) and an equal volume of fresh culture medium for infection. Mouse fibroblast and pre-B cell reprogramming was conducted by plating 5,000 cells per well in 6-well gelatine-coated plates in N2B27 2i/LIF or bFGF/activin conditions supplemented with doxycyclin. shRNA and overexpression experiments were conducted on $Utx^{Δ/Y}$ or $Utx^{+/Y}$ MEF cells as indicated, which were subjected to lentiviral shRNA infection or Amaxa transfection (Lonza) with mammalian expression vectors. The following lentiviral vectors (pLKO.1-puro) purchased from Sigma Aldrich were used: murine Utx shRNA no. 1, TRCN0000096243; human UTX shRNA no. 1, TRCN0000107760; human UTX shRNA no. 2, TRCN0000107761; mouse Jmjd3 shRNA no. 1, TRCN0000095266; mouse Jmjd3 shRNA no. 2, TRCN0000095268; human JMJD3 shRNA no. 1, TRCN0000236677; human JMJD3 shRNA no. 2, TRCN0000236678. The following shRNAs were obtained from Open Biosystems: Sall4, RMM4534-NM_201396; Sall1, RMM4534-NM_021390; Utf1, RMM4534-NM_009482. pLKO.1-puro empty vector control or scrambled hairpin containing plasmid DNA were used as controls. Eed siRNA (ON-TARGETplus SMARTpool L-049898-00) and the control siRNA (ON-TARGETplus Non-targeting pool

D-001810-10-05) were purchased from Dharmacon. Ten nanomoles siRNA or control was used for each transfection of NGFP1 $Utx^{Δ/Y}$ MEFs already on doxycycline for 7 days with lipofectamine RNAiMAX (Invitrogen). Embryonic germ cells were derived from E8.5 mouse chimaeric embryos as previously described[22].

DNA plasmids. The following mammalian constitutive overexpression vectors were used in somatic cells undergoing reprogramming following Amaxa Nucleoeffector transfection (Lonza): pCS2-UTY-F, PCS2-UTX-F, pCS2-JMJD3-F, pCMV-HA-JMJD3 (Addgene). Constitutively expressed lentiviruses FUW-Utx, FUW-UTX and FUW-UTXH1146A were generated by insert cloning into EcoRI sites of the FUW vector to generate constitutive expression following viral transduction and stable integration in somatic or PSC lines. pMXs moloney viruses encoding Myc, Nanog, Sall4 and Utf1 were obtained through Addgene.

Immunoprecipitation and immunoblotting analyses. HEK293T cells were transfected with each cDNA clone in an expression vector using jetPEI (Polyplus transfection) and were lysed 48 h later in lysis buffer (5 mM Tris-HCl pH 7.4, 15 mM NaCl, 0.1% Triton and 1.5 mM EDTA). The following plasmids were used for transfections in different combinations: pCMV-HA-UTX, pCMV-HA-JMJD3 pMSCV-Flag-OCT4, pMSCV-Flag-SOX2, pMSCV-Flag-KLF4 (obtained through Addgene). Thirty microlitres of anti-FlagM2 beads (Sigma) were incubated for 4 h in the anti-Flag fractions, where 5 μg of IgG were added to the anti-IgG fractions for 4 h and then 30 μl of G-sepharose beads were added for another 4 h. Subsequently, both fractions (the anti-Flag and anti-IgG) were centrifuged and the beads were washed six times with lysis buffer. The binding proteins were eluted with 1 mg ml^{-1} of Flag peptide buffer for the anti-FlagM2 beads (except for OCT4 where the beads where boiled) or by boiling with sample buffer, and analysed by SDS–PAGE and immunoblotting. The immunoblot analyses were performed using the following antibodies: anti-Flag (clone M2, F3165, Sigma), anti-Flag (M2 affinity gel, A2220, Sigma), anti-HA (clone 16B12 MMS-101R, Covance).

Mouse embryo micromanipulation. Pluripotent stem cells (ES cells or iPSCs) were injected into BDF2 diploid blastocysts. Microinjection into blastocysts placed in M16 medium under mineral oil was done by a flat-tip microinjection pipette. A controlled number of 10–12 cells were injected into the blastocyst cavity. After injection, blastocysts were returned to KSOM media (Invitrogen) and placed at 37 °C until transferred to recipient females. Ten to fifteen injected blastocysts were transferred to each uterine horn of 2.5 days post-coitum pseudo-pregnant females. 4n tetraploid complementation assay was performed by fusing BDF2 embryos at the two-cell stage, and subsequently allowing the embryos to develop until the blastocyst stage at day 3.5, and were then used for PSC micro-injection. Embryos were recovered for analysis at different time points during development or allowed to develop to full term. Determining germline transmission was performed by mating chimaeric animals with C57B/6 females, and continuous checking for agouti-coloured pups. At least four separate litters from each male chimaera had to be obtained before determining lack of contribution to the germ line. All animal studies were conducted according to the guideline and following approval by the Weizmann Institute Institutional Animal Care and Use Committee (approval no. 00960212-3).

***In vitro* PGC differentiation.** BVSC 40XY mouse ES cell line[21] carrying Blimp1–mVenus and Stella–CFP transgenes was maintained in N2B27 2i/LIF naive growth conditions. BVSC $Utx^{+/Y}$ (or $Utx^{flox/Y}$), and $Utx^{Δ/Y}$ ES cells were differentiated in PGC-like cells *in vitro* as previously described[21]. For embryonic germ cell experiments, 20,000 cells expressing both Blimp1–GFP and Stella–CFP were plated in 6-well plates in N2B27 2i/LIF medium.

Immunoflourescence staining. For immunoflourescence of gonads, embryos from blastocyst injection of $Utx^{flox/Y}$ BVSC, Utx$^{Δ/Y}$ BVSC and $Utx^{Δ/Y}$ V6.5 ES cells were dissected at ~E12.5 or at E10.75. Embryos were further dissected at the forelimb level and the anterior half was stained for X-Gal to determine the chimaeric embryos. The posterior half or isolated gonads (dissected from the posterior portion with attached mesonephros) of E10.75 and E12.5 embryos, respectively, were placed in mesh inserts (Corning, 3477), fixed in 4% PFA in 0.1 M PB (pH 7.2) at 4 °C for 2–3 h, washed three times at 4 °C in PB for 10 min, cryoprotected by sequential incubation in 10, 20 and 30% sucrose (each for at least 6 h) in PB at 4 °C, embedded in Tissue-Tek OCT compound (Sakura), frozen on dry ice, and stored at −70 °C. Sections (10–12 μm) were cut, mounted on Superfrost plus slides (Thermo Scientific), and stored up to 6 months at −70 °C. For immunoflurecence, sections were equilibrated to room temperature (25 °C), rinsed in PBS, and blocked and permeabilized in Blocking solution (5% normal donkey serum, 0.5% BSA, 0.5% Glycine, 0.1% Triton X-100 in PBS) for 1 h at room temperature (for epigenetic markers staining, sections were rinsed in 0.5% Triton/PBS before blocking step to improve permeabilization). Slides were then incubated in the appropriate primary antibody diluted in blocking solution at 4 °C overnight. Sections were then washed three times (5 min each) in PBS plus 0.1%

Triton X-100, incubated with appropriate secondary antibodies at room temperature for 1 h, counter stained with DAPI (1 µg ml^{-1}, 0215754; MP Biomedicals) and mounted with Shandon Immu-Mount (Thermo Scientific, 9990412). All images were collected on a Zeiss Axioscope microscope and LSM710 confocal microscope and processed with ImageJ software (NIH) and Adobe Photoshop CS4. The following antibodies were used: rabbit anti-MVH (1:400, ab13840; Abcam), rabbit anti-Nanog (1:200, A300-397A; Bethyl), mouse anti-Oct4 (1:100, 611203; BD), chick anti-GFP (1:2,000, ab13970; Abcam), chick anti-βgal (1:1,000, ab9361;Abcam), rabbit anti-Sall4 (1:500, ab29112; Abcam), rabbit anti-H3K27me3 (1:500/1:200 for E12.5 gonad/E10.75 embryo section, 07-449; Upstate), rabbit anti-H3K4me3 (1:400/1:200 for E12.5 gonad/E10.75 embryo sections, Ab8580; Abcam), rabbit anti-H3K4me2 (1:400/1:200 for E12.5 gonad/E10.75 embryo sections, Ab32356; Abcam), rabbit anti-H3K9me3 (1:500/1:200 for E12.5 gonads/E10.75 embryo sections, Ab8898; Abcam).

RT–PCR analysis. Total RNA was isolated using the RNeasy Kit (Qiagen). Three micrograms of total RNA was treated with DNase I to remove potential contamination of genomic DNA using a DNA Free RNA kit (Zymo Research). One microgram of DNase-I-treated RNA was reverse transcribed using a First Strand Synthesis kit (Invitrogen) and ultimately re-suspended in 100 µl of water. Quantitative PCR analysis was performed in triplicate using 1/50 of the reverse transcription reaction in an Viia7 platform (Applied Biosystems). Error bars indicate standard deviation of triplicate measurements for each measurement.

Gene expression data acquisition, processing and analysis. Total RNA was isolated from indicated cell lines. The concentration of RNA was quantified and subjected to quality control on Agilent Bioanalyzer. Two-hundred and fifty nanograms of RNA were simultaneously processed from each sample. cDNA was fragmented, labelled and hybridized to Affymetrix Mouse Gene 1.0 ST GeneChip (Affymetrix), which contained 35,557 probes. Transcript levels were processed from image files using the RMA method[23], which corrects for non-biological sample variation using quantile normalization, implemented by the Affymetrix Expression Console software. Data were further filtered to include probes that have at least one call higher than 64 ($= 2^6$), resulting in 21,811 probes that are mapped to 15,815 RefSeq transcripts (http://www.ncbi.nlm.nih.gov/RefSeq/). For gene expression statistical analysis, we used Cluster software (http://rana.lbl.gov/EisenSoftware.htm) to run hierarchical clustering on the samples, using complete linkage centred correlation as a distance metric. We used Matlab version R2011a to run principle components analysis that detects the principle components with the largest variation in the data. Differentially expressed gene signatures include transcripts that are changed more than twofold in $Utx^{\Delta/Y}$ compared to $Utx^{+/Y}$, or in MEFs with doxycycline compared to MEFs without doxycycline. For the functional analysis the signature genes were mapped from RefSeq transcripts to Entrez Gene accessions.

Functional annotation analysis. To annotate gene signatures we tested for their enrichment in functional gene sets. Functional gene sets were taken from the Gene Ontology (GO; http://www.geneontology.org), KEGG (http://www.kegg.com) and Biocarta (http://www.biocarta.com) databases. Protein–DNA binding annotations were taken from various publications[20,24–29]. Importantly, two of the data sets used contained fully genetically identical donor cells (on V6.5 ES cell C57B/6 X 129 F1 genetic background), whereas others were conducted on a highly related pure 129 mouse genetic background or on human embryonic cell lines. In the latter cases, human genes were mapped to mouse genes using Homologene database (http://www.ncbi.nlm.nih.gov/homologene). Data were taken from original publications without further post-processing, as such analysis indentifies global trends and largely overcomes noisy signatures. Enrichment P values were calculated using Fisher exact test and corrected for multiple hypotheses using a FDR threshold of 1%.

Chromatin immunoprecipitation and sequencing library preparation. Approximately 60×10^6 cells were cross-linked in formaldehyde (1% final concentration, 10 min at room temperature), and then quenched with glycine (5 min at room temperature). Fixed cells were lysed in 50 mM HEPES KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40 alternative, 0.25% Triton supplemented with protease inhibitor at 4 °C (Roche, 04693159001), centrifuged at 950g for 10 min and re-suspended in 0.2% SDS, 10 mM EDTA, 140 mM NaCl and 10 mM Tris-HCl. Cells were then fragmented with a Branson Sonifier (model S-450D) at −4 °C to a size range between 200–800 bp, and precipitated by centrifugation. Ten micrograms of each antibody were pre-bound by incubating with Protein-G Dynabeads (Invitrogen100-07D) in blocking buffer (PBS supplemented with 0.5% Tween and 0.5% BSA) for 2 h at room temperature. Washed beads were added to the chromatin lysate, and then incubated overnight. Samples were washed 5 times with RIPA buffer, twice with RIPA buffer supplemented with 500 mM NaCl, twice with LiCl buffer (10 mM TE, 250 mM LiCl, 0.5% NP-40, 0.5% DOC), once with TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA), and then eluted in 0.5% SDS, 300 mM NaCl, 5 mM EDTA, 10 mM Tris-HCl pH 8.0 at 65 °C. Eluate was incubated in 65 °C for 4 h, and then treated sequentially with RNaseA (Roche, 11119915001) for 30 min and Proteinase K for 2 h. DNA was purified with the Agencourt AMPure XP system (Beckman Coulter). Libraries of cross-reversed ChIP DNA samples were prepared according to a modified version of the Illumina Genomic DNA protocol, as described previously[30]. The following antibodies were used for chromatin immunoprecipitation experiments: control IgG (ChIP grade, ab46540; Abcam), anti-histone H3 trimethyl K4 (ChIP grade, ab8580; Abcam), anti-histone H3 trimethyl K27 (ChIP grade, 07-449; Millipore), anti-KDM6A Utx (ab84190; Abcam) and affinity-purified anti-Utx serum as previously described[31]. Throughout the manuscript, genomic localization and gene expression experiments, doxycycline treatment was conducted for 8 days.

ChIP-seq data analysis. We used bowtie software[32] version 0.12.5 to align reads to the mouse mm9 reference genome (University of California, Santa Cruz (UCSC), July 2007). We only considered reads that were uniquely aligned to the genome with up to a single mismatch, taking the single best match of each read. We identified enriched intervals of H3K4me3, H3K27me3 and Utx using MACS version 1.4.1[33]. We used sequencing of whole cell extract as a control in order to define a background model. Duplicate reads aligned to the exact same location are excluded by MACS default configuration. Enriched intervals were mapped to genes if they overlapped a 3 Kb symmetric interval before and after their TSSs (taken from the UCSC mm9 Known Gene Table). Functional enrichment was calculated in the same way as for gene expression (described above). Reads landscape was processed and visualized using IGV software version 2.0[34]. ChIP-seq data on wild-type samples were highly reproducible in comparison to previous publications[2,25] (data not shown). For motif detection, motifs that are enriched in Utx binding locations were detected using the SeqPos tool[35] in the Cistrome package (http://cistrome.org/ap/). Utx peaks in ES cells were run against the Cistrome-curated motif database. For Utx coverage analysis, the mean density of Utx binding reads was calculated for each gene in the UCSC Known Genes Table (version mm9). The fold change between the mean density in MEFs with doxycycline (8 days) and MEFs without doxycycline (day 0) was then calculated for each gene. Binding profiles around the TSS were calculated using SitePro software in the Cistrome package (http://cistrome.org/ap/). The profiles over all mouse genes were calculated using the UCSC Known Genes Table (version mm9), including 55,419 regions that are mapped to 29,952 Entrez genes. The average number of reads in a 50-bp window was calculated 3 Kb before and after the TSS. These averages were then converted to a Z-score by normalizing each window with the mean and standard deviation of that window, as was calculated for each sample and over 10,000 random locations in the genome. The profiles over OSKMN targets were calculated in a similar manner, but over 805 genes that were previously found to be bound by at least 3 of the following factors: Oct4, Sox2, Klf4, Myc and Nanog[20]. The profiles around Utx binding peaks were calculated in a similar manner, but only over the peaks identified by MACS software. Chromatin heatmap was generated using the SeqMINER program directly from aligned read files, without further normalization.

22. Leitch, H. G. et al. Embryonic germ cells from mice and rats exhibit properties consistent with a generic pluripotent ground state. Development 137, 2279–2287 (2010).
23. Irizarry, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249–264 (2003).
24. Boyer, L. A. et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947–956 (2005).
25. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448, 553–560 (2007).
26. Loh, Y. H. et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nature Genet. 38, 431–440 (2006).
27. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132, 1049–1061 (2008).
28. Bruce, A. W. et al. Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. Proc. Natl Acad. Sci. USA 101, 10458–10463 (2004).
29. Zhao, X. D. et al. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 1, 286–298 (2007).
30. Mikkelsen, T. S. et al. Dissecting direct reprogramming through integrative genomic analysis. Nature 454, 49–55 (2008).
31. De Santa, F. et al. The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell 130, 1083–1094 (2007).
32. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25 (2009).
33. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
34. Robinson, J. T. et al. Integrative genomics viewer. Nature Biotechnol. 29, 24–26 (2011).
35. He, H. H. et al. Nucleosome dynamics define transcriptional enhancers. Nature Genet. 42, 343–347 (2010).

LETTER

The dynamic disulphide relay of quiescin sulphydryl oxidase

Assaf Alon[1], Iris Grossman[1], Yair Gat[1], Vamsi K. Kodali[2], Frank DiMaio[3], Tevie Mehlman[4], Gilad Haran[5], David Baker[3], Colin Thorpe[2] & Deborah Fass[1]

Protein stability, assembly, localization and regulation often depend on the formation of disulphide crosslinks between cysteine side chains. Enzymes known as sulphydryl oxidases catalyse *de novo* disulphide formation and initiate intra- and intermolecular dithiol/ disulphide relays to deliver the disulphides to substrate proteins[1,2]. Quiescin sulphydryl oxidase (QSOX) is a unique, multi-domain disulphide catalyst that is localized primarily to the Golgi apparatus and secreted fluids[3] and has attracted attention owing to its over-production in tumours[4,5]. In addition to its physiological importance, QSOX is a mechanistically intriguing enzyme, encompassing functions typically carried out by a series of proteins in other disulphide-formation pathways. How disulphides are relayed through the multiple redox-active sites of QSOX and whether there is a functional benefit to concatenating these sites on a single polypeptide are open questions. Here we present the first crystal structure of an intact QSOX enzyme, derived from a trypanosome parasite. Notably, sequential sites in the disulphide relay were found more than 40 Å apart in this structure, too far for direct disulphide transfer. To resolve this puzzle, we trapped and crystallized an intermediate in the disulphide hand-off, which showed a 165° domain rotation relative to the original structure, bringing the two active sites within disulphide-bonding distance. The comparable structure of a mammalian QSOX enzyme, also presented here, shows further biochemical features that facilitate disulphide transfer in metazoan orthologues. Finally, we quantified the contribution of concatenation to QSOX activity, providing general lessons for the understanding of multi-domain enzymes and the design of new catalytic relays.

The introduction of disulphide bonds into folding proteins is typically accomplished by enzyme pairs[1,2]; one enzyme generates disulphides *de novo*, often with the aid of a bound cofactor, and the second enzyme acquires these disulphides by dithiol/disulphide exchange and passes them on to substrate proteins. Examples of such partnerships include DsbB and DsbA in the *Escherichia coli* periplasm and, in eukaryotes, Ero1 and protein disulphide isomerase (PDI) in the endoplasmic reticulum and Erv1 and Mia40 in the mitochondrial intermembrane space. A partnership between vitamin K epoxide reductase (VKOR) and membrane-anchored PDI-family proteins has also been identified in the endoplasmic reticulum[6]. The QSOX enzymes[3] are the only known example of conserved concatenation of disulphide-generating and disulphide-transferring modules within a single polypeptide (Fig. 1a), although a VKOR-like protein is fused to its partner in a few species[7]. Domain fusion may allow QSOX to function under dilute conditions in extracellular environments[8–10]; other disulphide-generating enzymes function in sequestered intracellular compartments with high local concentrations of disulphide-transferring enzymes.

Formation and transfer of disulphide bonds in QSOX are mediated by redox-active cysteine pairs in the canonical pattern Cys-X-X-Cys

(Fig. 1a). One such dicysteine motif is in a thioredoxin-fold domain (Trx1) related to the redox-active domains of PDI. Another is in a flavin adenine dinucleotide (FAD)-binding domain (Erv) related to the mitochondrial enzyme Erv1. The proposed QSOX mechanism[11,12] involves the generation of disulphides catalytically at the FAD-proximal dicysteine motif, intramolecular disulphide transfer to the Trx1 domain and finally intermolecular dithiol/disulphide exchange with the substrate (Fig. 1a). How the Trx1 redox-active residues interact alternately with the Erv domain and with substrate proteins remains unclear.

The *Trypanosoma brucei* parasite, the cause of African trypanosomiases, encodes a QSOX enzyme (TbQSOX)[13]. We determined the X-ray crystal structure of TbQSOX to 2.3 Å resolution (Supplementary Table 1). In the crystallized configuration, the Trx1 redox-active site faces outwards (Fig. 1b), where it would be readily accessible to nucleophilic attack by substrate cysteine thiolates. However, in this position, the Trx1 active site is 42 Å away from the Erv active site and would thus be unable, after disengaging from the substrate, to accomplish the next step in the proposed relay mechanism (Fig. 1a).

To test whether the distance between redox-active sites in the TbQSOX crystal structure is representative of their average separation in solution, we used fluorescence resonance energy transfer (FRET) spectroscopy. TbQSOX Cys 72 was mutated to leave one cysteine (Cys 69) free within the Trx1 active site for labelling with the donor fluorophore Pacific Blue C_5-maleimide. The bound FAD, although not fluorescent, served as an energy acceptor at the Erv active site, and the decrease in donor fluorescence was taken to reflect FRET efficiency. The use of an intrinsic chromophore as one member of the FRET pair minimized potential perturbations due to labelling. An average distance of 55 Å was measured between the label at position 69 and FAD (Supplementary Table 2), supporting a requirement for major conformational changes during the QSOX disulphide relay. Donor labels introduced at a variety of other positions across the enzyme surface provided a set of ensemble-averaged distances characterizing the resting state of the enzyme (Fig. 1c, Supplementary Figs 1–3 and Supplementary Tables II and IIIa–c).

To obtain structural insight into the intramolecular disulphide-transfer step (see Supplementary Discussion), we eliminated the second cysteine of both TbQSOX redox-active Cys-X-X-Cys motifs. The resulting mutant is able to form, but not resolve, the interdomain disulphide essential for catalysis. We purified (Supplementary Fig. 4) and crystallized the closed, interdomain disulphide-bonded version of TbQSOX (TbQSOX$_C$) and solved its structure to 3.3 Å resolution. When compared with the wild-type enzyme, a ~165° rotation of the Trx1 domain and associated helix brings the two redox-active sites within covalent-bonding distance in TbQSOX$_C$ (Fig. 1d and Supplementary Movie 1). Consistent with this closed configuration being a transient state during enzyme turnover, the surface complementarity between the TbQSOX$_C$ redox-active domains is greater than

[1]Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel. [2]Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716, USA. [3]Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA. [4]Department of Biological Research Support, Weizmann Institute of Science, Rehovot 76100, Israel. [5]Department of Chemical Physics, Weizmann Institute of Science, Rehovot 76100, Israel.
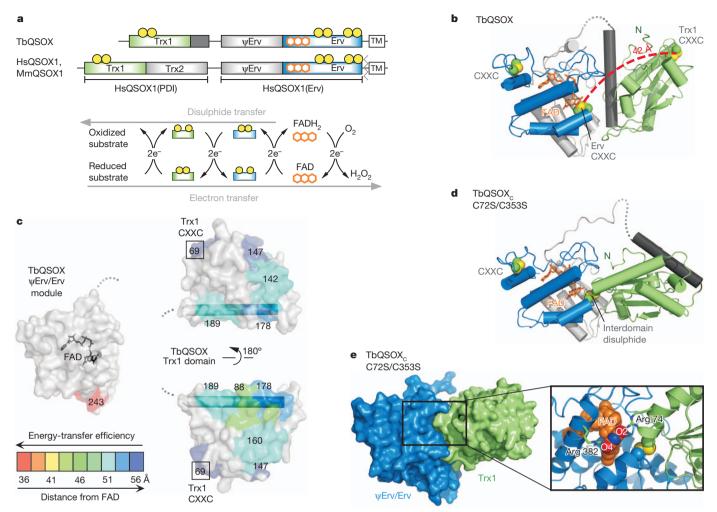
Figure 1 | TbQSOX undergoes domain reorientation to accomplish disulphide relay. a, Cys-X-X-Cys motifs are illustrated as pairs of yellow balls in maps of trypanosome and mammalian QSOX enzymes. Other cysteines are shown in Supplementary Fig. 13. Fused hexagons represent the FAD cofactor. A degenerate Erv-like domain[17] is designated 'ψErv'. Arrows depict the electron-transfer relay from reduced substrates to molecular oxygen and the corresponding outward flow of disulphide equivalents from QSOX to its substrates. TM, transmembrane. **b**, TbQSOX structure coloured according to

a. Disulphides in Cys-X-X-Cys motifs are in space-filling representation (Cβ atom, green; sulphur, yellow). Structure stereo views are depicted in Supplementary Fig. 14. **c**, Colour-coded intramolecular distances between fluorophore-labelled TbQSOX cysteines and the bound FAD cofactor. **d**, Structure of TbQSOX$_C$. **e**, Surface representation of TbQSOX$_C$ and zoom into the interdomain interface. FAD is in space-filling representation, and oxygen atoms in the isoalloxazine ring are coloured red and labelled. Structure figures were made using PyMOL (http://www.pymol.org).
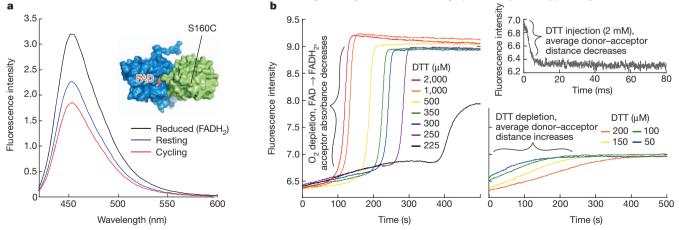
Figure 2 | Changes in the TbQSOX conformational ensemble during catalysis. a, Fluorescence of Pacific Blue conjugated at position 160 in resting TbQSOX (blue), TbQSOX oxidizing the model substrate dithiothreitol (DTT; red) and after oxygen depletion and conversion of the energy acceptor FAD to FADH$_2$ (black). Inset, surface representation of TbQSOX$_C$, as in Fig. 1e, with FAD and donor-labelling site indicated. **b**, Kinetics of donor fluorescence during TbQSOX oxidation of DTT. On DTT injection, labelled enzyme rapidly

converts to a state showing greater FRET efficiency (top right panel). Other panels show longer timescales at various initial DTT concentrations. At starting DTT concentrations above ~200 μM (left panel), oxygen becomes limiting, the flavin becomes trapped as FADH$_2$ (see Fig. 1a), flavin absorbance at ~450 nm drops markedly and donor fluorescence increases correspondingly. At starting DTT concentrations of 200 μM or below (bottom right panel), DTT is limiting and the enzyme returns to its oxidized, resting state.

non-evolved interfaces (represented by crystal contacts) and less than stable protein–protein interactions[14].

In addition to illustrating the functionally important reorientation of the redox-active sites, the TbQSOX$_C$ structure shows how closure of the Trx1 domain over the FAD-binding site may enhance the active-site chemistry for disulphide formation. In the reduced, anionic form of flavin, negative charge is distributed over atoms N1, O2 and O4 of the isoalloxazine ring. Other disulphide-forming enzymes position a basic residue near the cofactor to stabilize this negative charge[15,16]. In TbQSOX$_C$, the Trx1 domain contributes an arginine side chain (Arg 74) as an interaction partner for the FAD O2 atom (Fig. 1e), which may promote formation of a thiol–flavin charge-transfer complex and, in turn, a reduced state of FAD. Indeed, the TbQSOX Arg74Ala mutant is substantially less active than wild type (Supplementary Fig. 5). Other basic residues within the Erv domain (that is, His 356 and Arg 382) may further modulate the electrostatic environment of the FAD.

A second FRET study was conducted, this time monitoring energy transfer kinetically in the presence of substrate. The TbQSOX mutant with a donor fluorophore at residue 160 (Supplementary Table IIIc) showed a reproducible FRET increase when substrate was added (Fig. 2a). As expected, when oxygen was not limiting, the FRET signal returned to resting values as the substrate was exhausted (Fig. 2b). These observations indicate that the labelled position is closer to FAD in cycling TbQSOX than in the resting, native ensemble, further supporting conformational changes during catalysis.

Although TbQSOX has conserved functional features characteristic of the QSOX family, it is one of the most divergent orthologues. TbQSOX lacks a domain (Trx2) found in metazoan QSOX enzymes (Fig. 1a), and its Trx1 redox-active motif, Cys-Gly-Ala-Cys, differs from the Cys-Gly-His-Cys common to other QSOX proteins and to PDI-family proteins. We therefore expanded our study to include the architecture and dynamics of mammalian QSOX enzymes.

The structure of *Homo sapiens* QSOX1 (HsQSOX1) was solved in two complementary fragments, HsQSOX1(PDI) and HsQSOX1(Erv)[17] (Fig. 1a), crystallized separately (Fig. 3a and Supplementary Fig. 6). The HsQSOX1 domains containing redox-active Cys-X-X-Cys motifs were found to be similar to their TbQSOX counterparts (Supplementary
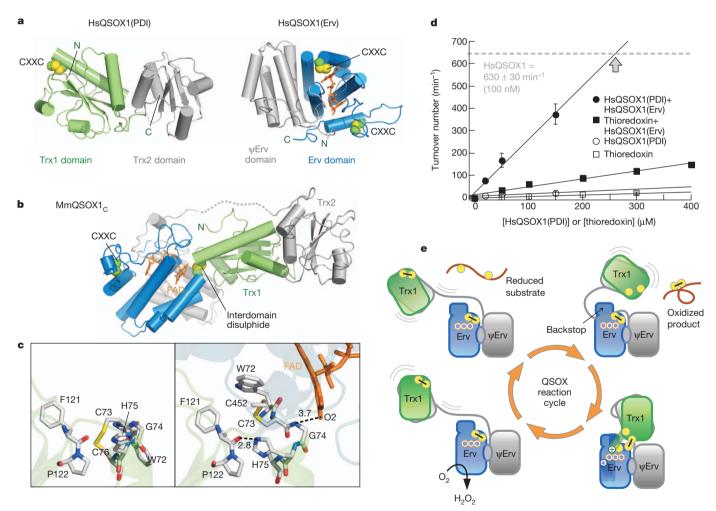

Figure 3 | Mammalian QSOX and mechanistic insights into the QSOX catalytic cycle. a, The structures of HsQSOX1(PDI) and HsQSOX1(Erv). Disulphides in Cys-X-X-Cys motifs are shown. Other cysteines appear in Supplementary Fig. 13. **b**, Structure of MmQSOX1$_C$. **c**, A comparison of HsQSOX1(PDI) (left) with MmQSOX1$_C$ (right) shows rearrangements in the Trx1 redox-active region on formation of the disulphide-transfer intermediate. Interatomic distances (dashed lines) are in Ångstrom. Residue numbering is according to MmQSOX1. **d**, Tethering increases the effective concentration of the HsQSOX1 disulphide-transferring module. Initial oxygen consumption rates were recorded for 100 nM HsQSOX1(Erv) and varying concentrations of HsQSOX1(PDI) (filled circles) or thioredoxin (filled squares) after injection of

DTT. Grey dashed line indicates the turnover number for intact HsQSOX1, measured at 100 nM. By extrapolation, ~250 μM HsQSOX1(PDI) with 100 nM HsQSOX1(Erv) would be expected to support a similar reaction rate (grey arrow). Error bars, s.d. **e**, Summary of the structural basis of the QSOX catalytic mechanism, as shown in this study. Adjacent yellow balls with black bar indicate disulphide bonds; separated yellow balls indicate reduced thiols or thiolates. The '+' symbols in the closed conformation (bottom right) represent arginine side chains from either the Trx1 domain (in TbQSOX) or the Erv domain (in TbQSOX and mammalian QSOX1) that may contribute to the electron-withdrawing ability of the FAD. The 'backstop' represents conserved Erv domain loops (Supplementary Fig. 9).

Table 4). Outside of these regions, the human and parasite structures are more divergent. In particular, HsQSOX1 has the Trx2 domain but lacks the long helix of TbQSOX (Fig. 1b). The two functional modules of HsQSOX1 seem to be tethered flexibly to one another, as found for TbQSOX (Supplementary Fig. 7 and Supplementary Table 5).

Mammalian QSOX mutants designed to mimic the interdomain disulphide-transfer intermediate, analogous to TbQSOX$_C$, were constructed. The *Mus musculus* structure, MmQSOX1$_C$, solved to 2.4 Å resolution, demonstrated how disulphide transfer readily accommodates the extra domain in metazoan QSOX (Fig. 3b). Moreover, some features that may orient the two enzyme halves are conserved (Supplementary Figs 8 and 9). One difference between TbQSOX$_C$ and MmQSOX1$_C$, however, is the presence of the histidine residue (His 75) in the MmQSOX1 Trx1 Cys-Gly-His-Cys motif. His 75 appeared in two distinct rotamers in the MmQSOX1$_C$ crystals, providing insight into the dynamics of catalysis (Supplementary Figs 10 and 11). In one configuration, His 75 would be in a position to interact transiently with the thiolate of the Erv active-site cysteine Cys 452 during resolution of the interdomain disulphide intermediate (Supplementary Fig. 11). A comparison of HsQSOX1(PDI) and MmQSOX1$_C$ shows how the end of the Trx1 redox-active helix rearranges on interaction with the FAD-binding domain (Fig. 3c). No Trx1 side chain comparable to TbQSOX Arg 74 interacts with the flavin in MmQSOX1$_C$, but unpaired -NH groups at the amino terminus of the MmQSOX1$_C$ Trx1 redox-active helix approach the FAD O2 atom (Fig. 3c). Furthermore, Arg 490 is recruited from within the Erv domain for interaction with FAD O4 (Supplementary Fig. 11), like its corresponding residue in TbQSOX (Arg 382). In summary, many structural and dynamic properties of mammalian and trypanosome QSOX proteins are similar, and both enzymes seem to recruit electron-withdrawing groups to modulate active-site electrostatics during catalysis. However, the mammalian QSOX crystal structures show for the first time, to our knowledge, the plasticity and sophistication of the Cys-Gly-His-Cys motif, present in numerous redox-active proteins of the PDI family[18].

The expansion of multi-domain protein architectures contributes to the growth of the protein universe[19] and the development of complex species[20]. Rarely is it possible, however, to determine quantitatively how domain fusion contributes to protein function. Owing to its modular structure, QSOX provides an opportunity to quantify the benefit of concatenation of functional units. We measured the extent of this benefit by adding increasing concentrations of HsQSOX1(PDI) to catalytic amounts of HsQSOX1(Erv) to determine the ratio necessary to achieve the reaction rate of intact HsQSOX1 (Fig. 3d). We observed that the effective concentration of the disulphide-transferring module is increased by more than 2,500-fold by its concatenation to the disulphide-generating module through a poorly conserved linker (Supplementary Fig. 12). Our results emphasize the principle, applicable to design and *in vitro* evolution of new catalytic or signalling relays, that the function of proteins acting in tandem may be greatly improved by concatenation using flexible generic linkers.

Although mammalian cells are replete with PDI-family proteins[18], which contain diverse numbers and arrangements of redox-active and -inactive domains, QSOX is the only enzyme with an oxidoreductase domain homologous to PDI fused to a domain that generates disulphide bonds *de novo*. QSOX is also exceptional in having a primary localization outside of the endoplasmic reticulum. The QSOX structures reported here show how the two functions of the enzyme, generation of disulphides and transfer to substrate proteins, are permitted through a loose tethering of the two modules (Fig. 3e). On one hand, the fusion is sufficient to greatly increase the effective concentration of the two active sites, but on the other it allows an enormous range of motion. Understanding how this flexibility may further contribute to interactions with substrate must await identification of the native targets of QSOX catalysis in the late secretory pathway and extra-cellular environment.

METHODS SUMMARY

Protein preparation. TbQSOX and its mutants were produced as described[13]. Enzymes were purified using nickel-nitrilotriacetic acid (Ni-NTA), digested with thrombin and reapplied to Ni-NTA. Unbound material was collected, concentrated and purified by gel filtration. HsQSOX1(PDI) and MmQSOX1 (Cys76Ala/Cys455Ser) were produced in BL21(DE3) pLysS *E. coli* and purified as for TbQSOX. For crystallization, HsQSOX1(PDI) was treated with a dimethylamine–borane complex[21]. TbQSOX (Cys72Ser/Cys353Ser) and MmQSOX1 (Cys76Ala/Cys455Ser) were incubated before gel filtration with maleimide-functionalized polyethylene glycol (PEG) to separate TbQSOX$_C$ and MmQSOX1$_C$ from thiol-containing species.

Protein crystallization. TbQSOX crystals were grown by vapour diffusion over a well solution containing 14–18% (w/v) 2 kDa PEG monomethyl ether and 10% ethylene glycol. Crystals of TbQSOX$_C$ were grown over 20% (w/v) 3.35 kDa PEG, 0.2 M sodium sulphate and 0.5% (w/v) polyvinylpyrrolidone K15. Crystals of methylated HsQSOX1(PDI) were grown over 20% (w/v) 3.35 kDa PEG and 0.2 M lithium sulphate monohydrate. Crystals of MmQSOX1$_C$ were grown over 0.1 M MES buffer, pH 6.5, and 12% (w/v) 20 kDa PEG (orthorhombic) or 0.1 M HEPES buffer, pH 7.0, 10% (v/v) 1,5-pentanediol and 5% (v/v) glycerol (monoclinic).

Structure solution. The TbQSOX structure was determined by molecular replacement using a partial human QSOX structure (Protein Data Bank (PDB) code 3LLK) and a TbQSOX Trx1 homology model. The TbQSOX$_C$ structure was determined by molecular replacement using wild-type TbQSOX fragments. The HsQSOX1(PDI) structure was determined as described[22] (search models, PDB codes 3ED3 and 2E0Q). The MmQSOX1$_C$ structure was determined by molecular replacement using HsQSOX1(PDI) and HsQSOX1(Erv).

FRET spectroscopy. Pacific Blue C$_5$-maleimide labelling was according to manufacturer's instructions. The excitation wavelength was 400 nm, and energy transfer was measured with reference to enzyme containing FADH$_2$. TbQSOX (Ser160Cys) wavelength scans in the presence of dithiothreitol (1 mM) were initiated within 10 s after mixing and completed within 1 min. TbQSOX (Ser160Cys) kinetic experiments were conducted using a stopped-flow fluorimeter with a 420 nm long-pass emission filter.

Full Methods and any associated references are available in the online version of the paper.

1. Riemer, J., Bulleid, N. & Herrmann, J. M. Disulfide formation in the ER and mitochondria: two solutions to a common process. *Science* **324**, 1284–1287 (2009).
2. Tu, B. P. & Weissman, J. S. Oxidative protein folding in eukaryotes: mechanisms and consequences. *J. Cell Biol.* **164**, 341–346 (2004).
3. Kodali, V. K. & Thorpe, C. Oxidative protein folding and the quiescin–sulfhydryl oxidase family of flavoproteins. *Antioxid. Redox Signal.* **13**, 1217–1230 (2010).
4. Antwi, K. *et al.* Analysis of the plasma peptidome from pancreas cancer patients connects a peptide to overexpression of the parent protein in tumors. *J. Proteome Res.* **8**, 4722–4731 (2009).
5. Song, H. *et al.* Loss of *Nkx3.1* leads to the activation of discrete downstream target genes during prostate tumorigenesis. *Oncogene* **28**, 3307–3319 (2009).
6. Schulman, S., Wang, B., Li, W. & Rapoport, T. A. Vitamin K epoxide reductase prefers ER membrane-anchored thioredoxin-like redox partners. *Proc. Natl Acad. Sci. USA* **107**, 15027–15032 (2010).
7. Goodstadt, L. & Ponting, C. P. Vitamin K epoxide redutase: homology, active site and catalytic mechanism. *Trends Biochem. Sci.* **29**, 289–292 (2004).
8. Zanata, S. M. *et al.* High levels of active quiescin Q6 sulfhydryl oxidase (QSOX) are selectively present in fetal serum. *Redox Rep.* **10**, 319–323 (2005).
9. Ostrowski, M. C. & Kistler, W. S. Properties of a flavoprotein sulfhydryl oxidase from rat seminal vesicle secretion. *Biochemistry* **19**, 2639–2645 (1980).
10. Jaje, J. *et al.* A flavin-dependent sulfhydryl oxidase in bovine milk. *Biochemistry* **46**, 13031–13040 (2007).
11. Raje, S. & Thorpe, C. Inter-domain redox communication in flavoenzymes of the quiescin/sulfhydryl oxidase family: role of a thioredoxin domain in disulfide bond formation. *Biochemistry* **42**, 4560–4568 (2003).
12. Heckler, E. J., Alon, A., Fass, D. & Thorpe, C. Human quiescin-sulfhydryl oxidase, QSOX1: probing internal redox steps by mutagenesis. *Biochemistry* **47**, 4955–4963 (2008).
13. Kodali, V. K. & Thorpe, C. Quiescin sulfhydryl oxidase from *Trypanosoma brucei*: catalytic activity and mechanism of a QSOX family member with a single thioredoxin domain. *Biochemistry* **49**, 2075–2085 (2010).
14. Mitra, P. & Pal, D. New measures for estimating surface complementarity and packing at protein–protein interfaces. *FEBS Lett.* **584**, 1163–1168 (2010).
15. Inaba, K., Takahashi, Y. H., Ito, K. & Hayashi, S. Critical role of a thiolate-quinone charge transfer complex and its adduct form in *de novo* disulfide bond generation by DsbB. *Proc. Natl Acad. Sci. USA* **103**, 287–292 (2006).
16. Fass, D. The Erv family of sulfhydryl oxidases. *Biochim. Biophys. Acta* **1783**, 557–566 (2008).

17. Alon, A., Heckler, E. J., Thorpe, C. & Fass, D. QSOX contains a pseudo-dimer of functional and degenerate sulfhydryl oxidase domains. *FEBS Lett.* **584,** 1521–1525 (2010).
18. Appenzeller-Herzog, C. & Ellgaard, L. The human PDI family: versatility packed into a single fold. *Biochim. Biophys. Acta* **1783,** 535–548 (2008).
19. Levitt, M. Nature of the protein universe. *Proc. Natl Acad. Sci. USA* **106,** 11079–11084 (2009).
20. Tordai, H., Nagy, A., Farkas, K., Bányai, L. & Patthy, L. Modules, multidomain proteins and organismic complexity. *FEBS J.* **272,** 5064–5078 (2005).
21. Walter, T. S. *et al.* Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14,** 1617–1622 (2006).
22. DiMaio, F. *et al.* Improved molecular replacement by density- and energy-guided protein structure optimization. *Nature* **473,** 540–543 (2011).

Author Contributions A.A. designed experiments, expressed, purified and crystallized proteins, and, together with D.F., solved and refined the TbQSOX, TbQSOX$_C$ and MmQSOX1$_C$ structures. A.A. also performed the FRET and crosslinking experiments. I.G. improved the TbQSOX crystals. Y.G. grew the MmQSOX1$_C$ crystals. V.K.K. and C.T. provided plasmids and helped to design and analyse experiments. F.D. and D.B. accomplished the molecular replacements to solve the HsQSOX1(PDI) structure. T.M. performed the mass spectrometry experiments and analyses. G.H. helped to design and analyse the FRET experiments and assisted with operation of the fluorimeter. D.F. expressed proteins, performed oxygen consumption measurements and designed and analysed the experiments. A.A. and D.F. wrote the manuscript.

Author Information Atomic coordinates and structure factors for the TbQSOX, TbQSOX$_C$ and HsQSOX1(PDI) structures have been deposited with the Protein Data Bank under accession codes 3QCP, 3QD9 and 3Q6O, respectively. The MmQSOX1$_C$ coordinates and structure factors have been deposited with accession codes 3T58 (orthorhombic) and 3T59 (monoclinic). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.F. (deborah.fass@weizmann.ac.il).

METHODS

Plasmid construction. TbQSOX mutants not previously available were made using the QuikChange mutagenesis kit (Stratagene) on the basis of the published TbQSOX expression plasmid[13]. HsQSOX1(PDI) spans residues 33–272 of HsQSOX1. The coding sequence for this region was cloned between the NdeI and BamHI sites of the pET-15b vector (Novagen). The HsQSOX1 protein used for activity assays spans residues 33–546 of HsQSOX1. The coding sequence for this region was cloned between the NdeI and BamHI sites of the pET-15b vector. The thrombin cleavage site and His$_6$ tag encoded by the vector were excised by NcoI and NdeI restriction and replaced with a His$_6$ tag only. The MmQSOX1 (Cys76Ala/Cys455Ser) double mutant was made by mutation of a *Mus musculus* synthetic QSOX1 gene codon optimized for protein production in *E. coli* (GenScript). The coding sequence for residues 36–550 was inserted into pET-15b between the NdeI and BamHI sites.

Protein preparation. TbQSOX and its mutants were produced in the Origami 2 (DE3) *E. coli* cell strain as described[13], except that bacteria were grown in Luria–Bertani media. The MmQSOX1 (Cys76Ala/Cys455Ser) double mutant was produced in the BL21 (DE3) pLysS *E. coli* strain. Cells were lysed by sonication in 20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl and 20 mM imidazole, supplemented with protease inhibitors (100 μg ml^{-1} phenylmethylsulphonyl fluoride (PMSF) and 1 μg ml^{-1} of leupeptin, aprotinin and pepstatin A). Crude extract was clarified by centrifugation at 40,000*g* for 1 h. The enzymes, containing N-terminal His$_6$ tags, were purified from the soluble fraction by nickel-nitrilotriacetic acid (Ni-NTA) chromatography (GE Healthcare). Eluted TbQSOX, TbQSOX mutants and the MmQSOX1 mutant were exchanged into 20 mM sodium phosphate buffer, pH 7.4, 100 mM NaCl and 20 mM imidazole using a PD-10 desalting column (GE Healthcare). Thrombin (10 U per mg protein) was added, and the cleavage reaction was incubated for 2 h at 25 °C. PMSF was added to a concentration of 1 mM to inhibit thrombin and the protein was reapplied to Ni-NTA resin. Unbound material was collected, concentrated and applied to a PD-10 desalting column pre-equilibrated with crystallization stock buffer (10 mM Tris, pH 8, and 100 mM NaCl) and then applied to a HiLoad 16/60 Superdex 75 gel filtration column (GE Healthcare) under the same buffer conditions. Alternatively, size-exclusion chromatography was performed in 20 mM sodium phosphate buffer, pH 7.5, and 200 mM NaCl. The TbQSOX (Cys72Ser/Cys353Ser) mutant was incubated before gel filtration with maleimide-functionalized 5 kDa polyethylene glycol (PEG). The MmQSOX1 (Cys76Ala/Cys455Ser) mutant was incubated with maleimide-functionalized 10 kDa PEG. The fraction lacking the Cys 69–Cys 350 disulphide bond (for TbQSOX) or the Cys 73–Cys 452 disulphide bond (for MmQSOX1) became doubly modified by PEG, which increased its hydrodynamic radius sufficiently to allow separation from TbQSOX$_C$ or MmQSOX1$_C$ (Supplementary Fig. 4). Owing to high yields of the MmQSOX1 mutant, the peak corresponding to the protein not modified by PEG was collected, reconcentrated and loaded again on the gel-filtration column to achieve complete separation from PEG-modified material.

HsQSOX1(PDI) and HsQSOX1(Erv) were produced in the BL21(DE3) pLysS *E. coli* strain and purified as for wild-type TbQSOX, except that the HsQSOX1 construct lacked a thrombin cleavage site and was therefore subjected to gel filtration after the first Ni-NTA column without the cleavage step. Before crystallization, HsQSOX1(PDI) was treated with a dimethylamine–borane complex to methylate lysines as described[21], and the protein was then separated from excess reagents by gel filtration in crystallization stock buffer.

Protein crystallization. Crystals were grown by hanging-drop vapour diffusion at 293K. Wild-type TbQSOX and TbQSOX (Ser160Cys) crystals were grown over a well solution containing 14–18% (w/v) 2 kDa PEG monomethyl ether (MME) and 10% (v/v) ethylene glycol. Crystals were transferred to a solution containing 20% (w/v) 2 kDa PEG MME and 20% (v/v) ethylene glycol, and then transferred to a 1:1 mixture of mineral and Paratone oil (Exxon) before flash freezing. Crystals of TbQSOX$_C$ were grown over a well solution containing 20% (w/v) 3.35 kDa PEG, 0.2 M sodium sulphate and 0.5% (w/v) polyvinylpyrrolidone K15. Crystals were transferred to a solution of 25% (v/v) glycerol, 0.2 M sodium sulphate and 20% (w/v) 3.35 kDa PEG and flash frozen. Crystals of the methylated HsQSOX1(PDI) were grown over a well solution containing 20% (w/v) 3.35 kDa PEG and 0.2 M lithium sulphate monohydrate. Crystals were transferred to a solution of 25% (v/v) glycerol, 20% (w/v) 3.35 kDa PEG and 0.2 M lithium sulphate monohydrate and flash frozen. Crystals of MmQSOX1$_C$ grown over a well solution containing 0.1 M MES buffer, pH 6.5, 12% (w/v) 20 kDa PEG were of the orthorhombic space group $P2_12_12_1$. These crystals were transferred to a solution of 0.1 M MES, pH 6.5, 12% (w/v) 20 kDa PEG and 25% (v/v) glycerol for freezing. Crystals of MmQSOX1$_C$ grown over a well solution containing 0.1 M HEPES buffer, pH 7.0, 10% (v/v) 1,5-pentanediol and 5% (v/v) glycerol were of the monoclinic space group $P2_1$. These crystals were transferred to a solution of 0.1 M HEPES, pH 7.0, 10% (v/v) 1,5-pentanediol and 25% (v/v) glycerol, and then transferred to a 1:1 mixture of mineral oil and Paratone oil for freezing.

Data collection. Diffraction data were collected at 100K. Data for wild-type TbQSOX crystals, HsQSOX1(PDI) and the monoclinic ($P2_1$) MmQSOX1$_C$ crystals were collected at a wavelength of 1.544 Å on a RU-H3R generator (Rigaku) equipped with a RaxisIV++ image plate system and Osmic mirrors. TbQSOX data were collected to 2.3 Å resolution from crystals of space group $P2_12_12$. Data for the HsQSOX1(PDI) were collected to 2.05 Å resolution from a crystal of space group $P2_12_12_1$. Data for monoclinic MmQSOX1$_C$ were collected to 2.8 Å resolution from a crystal of space group $P2_1$. Data from the TbQSOX$_C$ crystals, of space group $P2_1$, were collected to 3.3 Å resolution at a wavelength of 0.9769 Å on European Synchrotron Radiation Facility (ESRF) beamline ID23-1 using a helical data-collection strategy (for rod-shaped crystals). Data for orthorhombic ($P2_12_12_1$) MmQSOX1$_C$ were collected to 2.4 Å resolution at a wavelength of 1.000 Å on ESRF beamline ID29. Data for TbQSOX (Ser160Cys) crystals were collected to 2.6 Å resolution at a wavelength of 1.005 Å on ESRF beamline BM30. Data for TbQSOX, TbQSOX (Ser160Cys), MmQSOX1$_C$ monoclinic and MmQSOX1$_C$ orthorhombic crystals were processed and scaled using DENZO and SCALEPACK[23]. TbQSOX$_C$ and HsQSOX1(PDI) data were processed and scaled using iMOSFLM[24] and SCALA[25].

Structure solution. The wild-type TbQSOX structure was determined by molecular replacement using Phaser[26]. The helical core of the partial human QSOX structure (Protein Data Bank (PDB) code 3LLK) and a homology model of the TbQSOX Trx1 domain generated using Modeller[27] were used as search models. The TbQSOX$_C$ structure was determined by molecular replacement using fragments of the wild-type TbQSOX structure as search models in Phaser. The HsQSOX1(PDI) structure was determined by molecular replacement, using search models from yeast Mpd1 (PDB code 3ED3) and K53E thioredoxin from *Sulfolobus tokodaii* (PDB code 2E0Q). A template based on 3ED3, trimmed of non-conserved loops and side chains, was first placed using Phaser; gaps in the template were rebuilt and refined using density- and energy-guided optimization with Rosetta[22]. This allowed placement of the second domain by Phaser using a trimmed model based on 2E0Q, which underwent the same rebuilding procedure. Model refinement was performed using CNS[28] or Phenix[29] incorporating the TLSMD[30] procedure. In all cases, model rebuilding was done using Coot[31]. The MmQSOX1$_C$ structure was determined using the $P2_12_12_1$ data and the structures of HsQSOX1(PDI) and HsQSOX1(Erv) as search models in Phaser. Chain A of the refined MmQSOX1$_C$ model was positioned (four copies) in the $P2_1$ MmQSOX1$_C$ crystal unit cell using Phaser and further refined using CNS[28]. As assessed using MolProbity[32], there are no Ramachandran outliers in the structures reported herein.

Surface complementarity. Normalized interface packing (NIP) and normalized surface complementarity (NSc) were calculated according to ref. 33. The interface area between the Trx1 domain and the ψErv/Erv domains in TbQSOX$_C$ was calculated to be 867 Å^2. The NIP and NSc values were calculated to be 2.33×10^{-4} and 3.87×10^{-4}, respectively. The interface area between the comparable domains of MmQSOX1$_C$ was calculated to be 892 Å^2. The NIP and NSc values were calculated to be 2.49×10^{-4} and 3.70×10^{-4}, respectively. The QSOX NIP values are greater than ~75% of the NIP values for evolved complexes and less than ~85% of the NIP values for crystal contacts of monomeric proteins. The QSOX NSc values are greater than ~80% of the NSc values for evolved complexes and less than 83% of the NSc values for crystal contacts. By comparison, the interface area of the two subunits in the Erv2 dimer (PDB code 1JR8) was calculated to be 1233 Å^2 and the NIP and NSc values were 1.81×10^{-4} and 2.30×10^{-4}, respectively. The Erv2 NIP value is greater than 39% of the NIP values for evolved complexes (that is, squarely within the expected packing values for complexes) and less than 95% of the NIP values for crystal contacts (that is, the packing is much better than most crystal contacts). The Erv2 NSc value is greater than 57% of the NSc values for evolved complexes (that is, the surface complementarity is typical of evolved complexes) and less than ~94% of the NSc values for crystal contacts (that is, the complementarity is much better than for most crystal contacts).

FRET spectroscopy. All mutants used in the FRET studies were purified as for the wild-type enzymes, except that a five- to tenfold molar excess of C$_5$ Pacific Blue-maleimide (Invitrogen) was added after gel filtration. A second gel filtration was then performed to remove excess dye. The fraction corresponding to the peak of the gel filtration chromatogram was used for further study. FRET experiments were performed in a 1 cm pathlength cuvette at a protein concentration of 6.3 nM in 20 mM sodium phosphate buffer, pH 7.5, 200 mM NaCl and 0.5 mM EDTA. Excitation was at a wavelength of 400 nm. As the fluorescence acceptor, namely the bound FAD cofactor, is an intrinsic part of QSOX enzymes, quantum yield of fluorophore conjugated at various positions in the native protein in the absence of acceptor could not be directly determined. Instead, acceptor absorbance was diminished *in situ* by reduction to FADH$_2$, which results in an 83% decrease in

absorbance at 454 nm. Reduction was accomplished by the addition of glucose oxidase and glucose to deplete the dissolved oxygen and DTT to reduce the TbQSOX active site as described in Supplementary Fig. 3. Donor fluorescence was monitored kinetically until an abrupt increase was observed, indicating conversion to $FADH_2$. Reduced state wavelength scans were then collected. Distances were calculated as described in Supplementary Table 2.

Mass spectrometry. Gel fragments were treated with varying protease combinations at 37 °C in 50 mM ammonium bicarbonate. Peptide mixtures were extracted from the gels with 80% acetonitrile and 1% trifluoroacetic acid (TFA), and the organic solvent was evaporated in a vacuum centrifuge. The resulting peptide mixtures were reconstituted in 80% formic acid and immediately diluted 1:10 with Milli-Q water before analysis. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) was performed using a 15 cm reversed-phase spraying fused-silica capillary column (inner diameter, 75 μm) made in-house and packed with 3 μm ReproSil-Pur $C_{18}AQ$ media (Dr. Maisch GmbH) using an UltiMate 3000 Capillary/Nano LC System (LC Packings, Dionex). The LC system was used in conjunction with an LTQ Orbitrap (Thermo Fisher Scientific) operated in the positive ion mode and equipped with a nanoelectrospray ion source. Peptides were separated with a 50-min gradient from 5 to 65% acetonitrile (buffer A, 5% acetonitrile, 0.1% formic acid and 0.005% TFA; buffer B, 90% acetonitrile, 0.2% formic acid and 0.005% TFA). The voltage applied to the union to produce an electrospray was 1.2 kV. The mass spectrometer was operated in the data-dependent mode. Survey mass spectrometry scans were acquired in the Orbitrap with the resolution set to a value of 60,000. Up to the six most intense ions per scan were fragmented and analysed in the linear trap. For the analysis of peptides, survey scans were recorded in the Fourier transform mode followed by data-dependent collision-induced dissociation of the six most intense ions in the linear ion trap. Raw data files were searched with MASCOT (Matrix Science) against a Swissprot database modified by inclusion of the relevant mutants. The program MassMatrix[34] was used for identification of disulfide bonds and BS3 crosslinks from MGF files.

QSOX enzyme assays. Sulphydryl oxidase activity was measured at 25 °C by monitoring oxygen consumption in a Clarke-type oxygen electrode (Hansatech Instruments Ltd). Buffer conditions were 50 mM potassium phosphate buffer, pH 7.5, 300 mM NaCl and 1 mM EDTA. HsQSOX1(PDI) and HsQSOX1(Erv) were mixed in the oxygen electrode chambe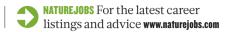r and reactions were initiated by injection of DTT to a concentration of 1 mM. For comparison, *E. coli* thioredoxin was prepared as described[35] and used in place of HsQSOX1(PDI). Intact HsQSOX1 activity was re-measured[12] for reference by diluting the enzyme to 100 nM into the oxygen electrode chamber and initiating the reaction by injection of DTT to 1 mM. Turnover numbers were calculated per FAD cofactor. Oxygen consumption by HsQSOX1(Erv) alone is indistinguishable from background rates under these conditions (not shown). Oxygen consumption in the presence of varying concentrations of HsQSOX1(PDI) alone or thioredoxin alone was minimal.
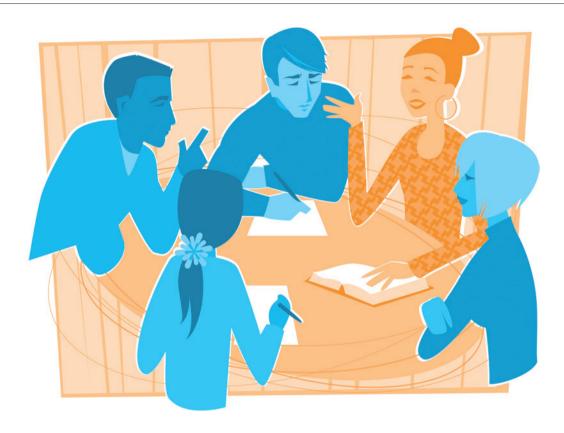
23. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
24. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newslett. Protein Crystallogr.* **no. 26** (1992).
25. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62,** 72–82 (2006).
26. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40,** 658–674 (2007).
27. Eswar, N., Eramian, D., Webb, B., Shen, M. Y. & Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **426,** 145–159 (2008).
28. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
29. Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. The Phenix refinement framework. *CCP4 Newslett.* **no. 42** (2005).
30. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* **62,** 439–450 (2006).
31. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
32. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66,** 12–21 (2010).
33. Mitra, P. & Pal, D. New measures for estimating surface complementarity and packing and protein–protein interfaces. *FEBS Lett.* **584,** 1163–1168 (2010).
34. Xu, H., Zhang, L. & Freitas, M. A. Identification and characterization of disulfide bonds in proteins and peptides from tandem MS data by use of the MassMatrix MS/MS search engine. *J. Proteome Res.* **7,** 138–144 (2008).
35. Gross, E. *et al.* Generating disulfides enzymatically: reaction products and electron acceptors of the endoplasmic reticulum thiol oxidase Ero1p. *Proc. Natl Acad. Sci. USA* **103,** 299–304 (2006).

CAREERS

R. COTTAM/IMAGEZOO/CORBIS

TRAINING

Workshops that work

Seminars on career alternatives and soft skills can provide crucial tips for advancement. But some workshops are more helpful than others.

BY TRISHA GURA

One Thursday afternoon in May, a conference room at the Beth Israel Deaconess Medical Center (BIDMC) in Boston, Massachusetts, is filling beyond capacity. More than two dozen postdocs and young faculty members from the BIDMC and other affiliates of Harvard Medical School in Boston stream in for tips on how to produce an effective oral presentation. The attendees, many of whom are not native English speakers, look overworked but expectant. They have relinquished the freedom of a late-afternoon coffee break because their career advancement is on the line. They know the importance of being able to deliver an hour-long lecture or a

ten-minute talk, daunting though that might be. "I have been giving talks for 30 years," begins the presenter, Terry Maratos-Flier, an endocrinologist and neurologist who directs the Office for Academic Careers and Faculty Development at the BIDMC. "So I figure I should take my expertise and offer it to you."

She does so in an hour, describing how to structure a talk, deliver it effectively and design clear, attractive slides. She engages her audience with a mixture of directives and humorous anecdotes; participants scribble notes, ask questions and offer their own accounts of success and catastrophe in public speaking.

"Sometimes I have a tendency to give more data when the audience has already got what they need," says Salin Dahlben, a psychiatrist

and neurologist at the BIDMC. "I try to add something more, perhaps because of my ego."

Maratos-Flier cautions that less is more, and suggests that speakers rehearse in front of colleagues, who can offer feedback such as 'Too much information' or 'You lost me there'. By the end of the talk, many participants feel much more confident in their presentation abilities.

As 'soft' communication skills become more important in research careers, early-career scientists are increasingly attending workshops to learn them. It is no longer enough to learn by watching and consulting supervisors. "When I was in training, we could figure this stuff out on our own, stumble around or hope we could get it from our mentors," says Maratos-Flier. ▶

▶ "Now, the sense is that individuals really need to have more directed education."

Graduate students and postdocs, under pressure to publish and seek funding, have little time to learn non-research skills to boost their competitiveness. But employers — especially in industry — want recruits to have those skills. "If you have waited until year five of your postdoc and you think you are going to break into industry as a 'plan B'," says Sibby Anderson-Thompkins, director of postdoctoral affairs at the University of North Carolina at Chapel Hill (UNC), "it may not happen."

So universities around the world are setting up schemes to teach transferable and soft skills. These might involve hour-long lectures, half-day seminars or hands-on, activity-based experiences that offer aspiring researchers advice and training on how to run a lab, create a budget, write a paper, collaborate with colleagues and, of course, deliver a good talk. But not all workshops provide useful or practical advice that is worth a fledgling scientist's time.

INFORMED DECISIONS

Early-career researchers looking for workshops should start on campus. Cathee Johnson Phillips, executive director of the US National Postdoctoral Association (NPA) in Washington DC, says that soft-skills training is now available at more than 135 offices or associations of postdoctoral affairs in the United States, as well as hundreds of career-development offices at universities and research institutions in the United States and Europe. The NPA itself organizes an annual meeting with workshops on a range of topics: everything from conflict resolution to English for postdocs. One seminar at this year's meeting, called 'Business Dining is Not About Food', taught etiquette for business lunches and dinners.

Once they have picked a provider, researchers must choose which workshops to attend. "It depends on what you want to get out of it," says Jill Hurst, a pharmaceutical and biochemical sciences postdoc at the UNC. "For students and younger postdocs trying to figure out which way they are going to go career-wise, a more formal, this-is-what-is-available panel discussion works well." For in-depth inquiries, it might be preferable to attend a lunch seminar with an industry representative, for example, says Hurst. The seminar format gives researchers a chance to ask probing questions and improve their sense of what day-to-day working lives are like in industry.

Some institutions offer more extensive experiences. Stephen Fuchs, a biochemist at Tufts University in Medford, Massachusetts, attended a leadership training course at Cold Spring Harbor Laboratory in New York. There, he and a select group spent three days listening, discussing and role-playing. "It was intense and focused," Fuchs recalls. "And everybody there was looking for the same types of things." Those included tips on negotiating salary and

position, managing personnel and dealing with difficult conversations in the lab.

The specificity made it the best training experience that Fuchs has ever had, he says. To assess just how focused a workshop will be, prospective attendees can contact organizers and find out what motivated them to put the workshop together. For example, if a company representative is giving a talk with the intention of recruiting scientists for certain positions, the event might not be a great choice for a first-year postdoc who simply wants to know more about working in industry. A panel discussion by industry experts would be a better fit.

Some misses can't be predicted. While still a postdoc at the UNC, Fuchs attended a seminar on the responsible conduct of research at the US National Institutes of Health (NIH) in Bethesda, Maryland. Speakers lectured on issues of authorship, data fabrication and conflict of interest. The seminar was mandatory for researchers who wanted to obtain federal funding, but it felt too "philosophical" to Fuchs, more a reflection on ethics than a guide to being a responsible researcher.

Scientists need practical guidance, says Fuchs. So he organized his own workshop on the subject at the UNC, and made sure that it included substantive advice. He brought in experts who deal with the issues on a daily basis, such as a representative from the UNC Office of Technology Transfer to talk about intellectual-property issues, and the university's ombudsman to discuss how he investigates complaints and mediates settlements. Participants gave positive feedback, and the workshop is now held regularly.

> "It is amazing, the lack of self-awareness I see."
>
> Sibby Anderson-Thompkins

Early-career scientists should keep an eye out for active rather than passive events, so that they can practise the skills of interest rather than merely discussing them. Dara Wilson-Grant, associate director of postdoctoral affairs at the UNC, held a workshop on networking, tied to an actual campus networking event with industry representatives. Workshop participants first listened to tips on networking, then role-played with each other and finally tried their luck at the event. Hurst attended and found the workshop so helpful that she landed a job interview that led to an offer after chatting with a science writer from a local biotechnology firm. "I learned how to ask good questions and get a sense of what the company was like," she says. "A year ago, I wouldn't have been able to do that."

Before attending a workshop, scientists can research the speaker's expertise and style

to ensure that they are a good fit. Dahlben's favourite thing about Maratos-Flier's workshop was her demeanour. By presenting herself as a seasoned researcher who had experienced the same travails as the audience and overcome them, Maratos-Flier had credibility with her audience, says Dahlben.

POCKETS OF RESISTANCE

Despite the need and opportunities for soft-skills training, not everyone is enthusiastic. Some see workshops as irrelevant or unnecessary. "In terms of giving a talk, I have participated in many conferences," says Bernard Charlier, a geosciences postdoc at the Massachusetts Institute of Technology in Cambridge. "I have seen very good presentations and very, very bad. So, I learn from those."

But it may not be possible to learn enough by imitation in today's competitive environment. Critiques and feedback are crucial. "It is amazing, the lack of self-awareness I see," says Anderson-Thompkins. "People think, 'This isn't something I need a lecture on'. But I have seen it in action, with networking, for example. Postdocs clear a room because of the way that they are approaching people."

Some early-career scientists might want to attend workshops, but run into resistance from supervisors who prefer their charges to stay in the lab rather than investigate alternative careers. However, workshop organizers such as Paule Biaudet, a specialist in doctoral studies and employability at Pierre and Marie Curie University in Paris, argue that professional development helps not only students and postdocs, but also their principal investigators, who should support their lab members' training aspirations. She explains to mentors that the reputation of a lab is often linked to the career successes of the people who have been trained there. "Then the supervisors feel they must let the students go to the training," says Biaudet.

Good training ultimately opens doors in an increasingly challenging job market. Maratos-Flier says that when she started out in research, "it would have been nice to have had somebody tell me in a non-judgemental way, 'you're doing this or that wrong', or a better or more effective way to do it. Now, you don't need to reinvent the wheel. Workshops can get you to a skill level you want to be at, without having to work it out all by yourself." ■

Trisha Gura *is a freelance writer based in Boston, Massachusetts.*

CORRECTION

The map in the story 'Excellence revisited' (*Nature* **487**, 519–521; 2012) omitted the symbol to show that the University of Konstanz has received an award for 'future concepts'. The map has been corrected online at go.nature.com/rciac4.

TURNING POINT
Nehal Mehta

Nehal Mehta is the inaugural Lasker Clinical Research Scholar — the first recipient of an award established by the US National Institutes of Health (NIH) and the Lasker Foundation in New York. Last month, Mehta moved from the University of Pennsylvania (UPenn) in Philadelphia to the NIH campus in Bethesda, Maryland, where he is studying links between cardiovascular disease and the chronic inflammation associated with conditions such as diabetes, obesity and psoriasis. The award affords a serious career boost: after 5–7 years with positive reviews, Mehta will have the option of getting tenure at the NIH or applying — with funding in tow — to an academic institution.

You trained as a medical doctor. When did you realize that you also wanted to do research?
It started with my internal-medicine residency. I had a lot of unanswered questions — for instance, why do certain people have heart attacks when they have atherosclerosis (hardening of the arteries) and others don't? My main mentor at UPenn encouraged me to investigate such questions. I enrolled about 80 people in his clinical study asking how inflammation affects molecular pathways that lead to atherosclerosis. He opened my eyes to clinical research.
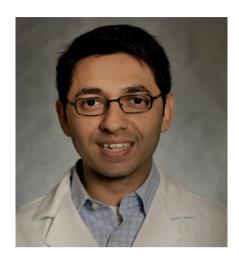
What benefits does the Lasker award bring?
It provides the resources to develop my career — not just money, but also crucial elements such as core facilities, access to patients and travel opportunities. The scholarship can last up to 12 years; clinical research is very time consuming, mostly because of clinical trials. This award comes once in a lifetime: I have the chance to run my own programme and see patients at the NIH Clinical Center, a world-class research hospital, without the barriers that face many young investigators, such as the need to secure funding.

What is your research focus?
My lab is poised to answer whether inflammation is simply a by-product of certain conditions, or a driver of cardiovascular disease. One of our studies asks whether treating chronic inflammation improves markers of cardiovascular health, such as how the body handles cholesterol and sugar.

What are your plans for the NIH?
I have a three-pronged approach. We are starting a national effort to obtain skin, fat and blood samples from people with psoriasis to understand the dysregulation of molecular and cellular pathways related to cardiovascular and metabolic diseases. The second arm will bring patients into the NIH Clinical Center to study risk factors for cardiometabolic disease and assess vascular inflammation. Finally, we will examine cardiometabolic pathways in animals and cell culture.

You plan to continue working at UPenn. How will you balance this with your job at the NIH?
I will be an adjunct faculty member at UPenn, and will go there every seven weeks to see people with psoriasis in my preventive cardiology programme. Hopefully most of them will come to the NIH; we will continue to enrol patients at UPenn but do assays and imaging analysis at the NIH. I have roots at UPenn: mentors are now collaborators. And I have an excellent nurse.

Did you ever consider getting a PhD?
I have thought about it, but I don't think it is absolutely necessary. There are some things I struggle with in the lab, such as operating certain equipment, but I can solve those problems by surrounding myself with great scientists. I do believe that a certain amount of advanced training is necessary. I have a master's degree in clinical epidemiology, which helps me to do high-level clinical research, design studies and work in a lab.

How do you juggle medicine and research?
My patients provide me with rich samples and rich questions. I get a lot out of each activity: by seeing patients, I am helping people but also doing research. ∎

INTERVIEW BY CHARLOTTE SCHUBERT

Industrial outcomes

Academic scientists funded by large companies publish, patent and are cited less often than those supported by smaller firms or other sources, a study says. *Research Grants, Sources of Ideas and the Effects on Academic Research*, released on 29 July by the Centre for European Economic Research in Mannheim, Germany, polled researchers at 46 German universities. Co-author Cornelia Lawson, an economist at the University of Turin, Italy, says that large grant-giving firms might modify the focus of the work during the project so that there are not enough data to publish after the grant. They also tend to fund applied research, which may be cited less often, she says. The study notes that industrial sponsorship has been on the rise globally for several years.

Awards for women

Five early-career female researchers from Africa, the Middle East, Asia, Latin America and the Caribbean are to be honoured with US$5,000 Elsevier Foundation Awards in a programme to encourage women to pursue science in nations that lack scientific expertise, resources and gender equality. "They are not given the opportunity to do good science," says Peter McGrath, a programme officer for co-sponsor TWAS, the academy of sciences for the developing world in Trieste, Italy. "This is a way to bring their work to the international forefront." Nominations for researchers from 81 eligible nations who earned PhDs in the past 10 years will be accepted until 30 September.

Initiative criticized

Germany's €4.6-billion (US$5.6-billion) Excellence Initiative has not helped universities to distinguish themselves, according to a briefing (go.nature.com/nsjpca) released on 22 June by the Social Science Research Center Berlin (WZB). The report says that institutions are concentrating on basic research to the detriment of teaching and applied research. "When we talk about excellence, we also have to talk about conditions for excellent teaching," says co-author Dagmar Simon, head of science-policy studies at the WZB. The scheme to make Germany's university system more competitive is in its second phase (see *Nature* **487**, 519–521; 2012).